

Hyaluronic Acid Production in *Streptococcus equi* Species

Robyn Milligan

Submitted for the degree of Doctor of Philosophy

Heriot-Watt University

Institute of Biological Chemistry, Biophysics and
Bioengineering

School of Engineering and Physical Sciences

December 2019

The copyright in this thesis is owned by the author. Any quotation from the thesis or use of any of the information contained in it must acknowledge this thesis as the source of the quotation or information.

Abstract

Hyaluronic acid (HA) is a glycosaminoglycan which exists ubiquitously in all organisms where it possesses a range of important functions. Due to the rheological properties and biocompatibility of HA solutions, it has been employed in a variety of medical applications. The rheology of a HA solution is a consequence of the concentration and molecular weight of the HA present. The current commonly used method for HA production is bacterial fermentation using Group C *Streptococcus* species which produce HA as a protective capsule in order to evade immune responses of their hosts. These are fastidious organisms and the fermentation process requires strict regulation of culture parameters in order to obtain optimum yields. A disadvantage of bacterial HA production is that the molecular weight of the product is often lower than the ideal range and the optimal yield is low at 5-10 g/L. The aim of this project is to investigate metabolic and genetic factors influencing HA capsule production in subspecies of the Group C *Streptococcus equi*. The study identifies a novel strain of *S. equi* capable of producing high concentrations of HA under the conditions studied and evaluates a currently used industrial procedure. Genomic and proteomic differences between subspecies of *S. equi* and mucoid/non-mucoid phenotypes have been investigated. The effects of the antibiotic phosphomycin on growth and HA production have also been explored. Collectively the results of this study have opened the scope of HA production research beyond the conventional methods used to date.

Acknowledgements

Firstly I would like to thank my supervisors. Professor Nicholas Willoughby for his support and encouragement throughout the entirety of my PhD, and Professor David Smith for his guidance, support and reassurance, for which I am so grateful.

I'd like to thank IB3 who provided a supportive and nurturing environment in which to carry out my project. I would especially like to thank PoBBBs members past and present, without whom my PhD would not have been the same. I have really enjoyed working and growing alongside such a great group of people, and have enjoyed the banter and fun in and outside of the university environment.

I can't express in words my gratitude to my husband Neil, for his patience, belief, support and love throughout my PhD. For never doubting me and for enduring the ups and downs with me.

I am thankful to my mum, dad and brother for their continued involvement, support and belief in me. I wouldn't be where I am today without you all. You'll be glad to know I think this is the last degree I'll collect.

Lastly I would like to thank my grandparents John and Betty McGibbon, who are sadly no longer with us. For their enduring love and support for me and their pride in all of my achievements in life. I wish you were both here and to you, I dedicate this thesis.

Table of Contents

1. Introduction	1
1.1. Polysaccharides	1
1.2. Hyaluronic Acid	2
1.2.1. Structure and Functions	2
1.2.2. History and production methods.....	5
1.2.3. Uses.....	6
1.2.3.1. Clinical.....	6
1.2.3.2. Tissue engineering	8
1.2.3.3. Moisturisers and cosmetics	8
1.3. Microbial Fermentation for HA production	9
1.4. Streptococci	11
1.4.1. Group C Streptococci	11
1.4.2. Hyaluronic Acid Capsule of Streptococci	13
1.4.2.1. Hyaluronic acid protection against phagocytosis	13
1.4.2.2. Genetic basis of capsule synthesis	14
1.4.2.3. Hyaluronan synthase.....	16
1.5. Project aims.....	17
2. Growth Parameters and Permutations in HA production by <i>S. equi</i> subspecies ...	18
2.1. Introduction.....	18
2.2. Materials and Methods	20
2.2.1. Culture media and Bacteria Strains.....	20
2.2.1.1. Fermentation media	20
Table 1. Fermentation media	21
Table 2. Sulphur Limited Broth.....	21
Table 3. Trace Salts Stock Solution	22
Table 4. Calcium Carbonate Stock Solution	22
Table 5. Sodium Molybdate Stock Solution.....	22
Table 6. Sulphur Limited Broth.....	22
2.2.1.2. Brain Heart Infusion Media and Agar plates.....	23
2.2.1.3. Bacteria Strains	23
2.2.2. Establishing growth in fermentation media in comparison to other media	23
2.2.2.1. Revival of Glycerol Stocks and Inoculum preparation.....	23
2.2.2.2. Multi-well plate growth curve set up.....	23
2.2.3. pH control using HEPES, PIPES and Tris buffers.....	24

2.2.4.	pH monitoring of <i>S. equi</i> cultures	24
2.2.5.	Capsule production analysis.....	25
2.3.	Results.....	25
2.3.1.	Colony morphology of SE40327 compared to other <i>S. equi</i> strains.....	25
2.3.2.	Growth variations between subspecies of <i>S. equi</i> in different media	28
2.3.3.	pH control and buffering.....	34
2.3.4.	Influence of media on capsule production of <i>S. equi</i> strains	41
2.4.	Discussion and Conclusions	43
2.4.1.	Sulphur limited media and its effect on HA production	43
2.4.2.	Selecting an appropriate buffer for pH control in small scale fermentation...45	
2.4.3.	Hyaluronic acid production in <i>S. equi</i> strains.....	48
3.	Comparative Genomics of <i>S. equi</i> subspecies	50
3.1.	Introduction.....	50
3.2.	Materials and Methods	51
3.2.1.	DNA extraction and sequencing.....	51
3.2.2.	Genome assembly, annotation, species comparison and MLST typing	52
3.2.3.	Analysis of hyaluronidase genes of each <i>S. equi</i> species	52
3.2.4.	Identification of hyaluronic acid associated genes in strains of interest	53
3.2.5.	Comparison of genes between strain phenotypes	53
3.2.6.	Comparison of <i>hasC</i> paralogs.....	53
3.3.	Results.....	54
3.3.1.	Hyaluronidase gene distribution across pathogenic Group C <i>Streptococci</i> species 54	
3.3.2.	Identification of hyaluronic acid associated genes.....	57
3.3.3.	Gene discrepancies between phenotypes.....	60
3.3.3.1.	Collagen-like surface protein	61
3.3.3.2.	M protein trans-acting positive regulator (Mga)	62
3.3.3.3.	RNA polymerase sigma-70 factor	63
3.3.4.	The <i>hasC</i> genes.....	64
3.3.5.	The MurA paralogs	66
3.3.5.1.	MurA1	66
3.3.5.2.	MurA2	68
3.4.	Discussion	71
3.4.1.	Hyaluronidases of <i>Streptococci</i>	71
3.4.1.1.	SE40327hyal1 and SE40327hyal2, true hyaluronidases?	74
3.4.1.2.	Beta-N-acetylglucosaminidases.....	77

3.4.2.	Identification of hyaluronic acid associated genes and the <i>has</i> operon	79
3.4.3.	Gene discrepancies between phenotypes.....	79
3.4.3.1.	Collagen-like surface protein	80
3.4.3.2.	M protein trans-acting positive regulator Mga	80
3.4.3.3.	RNA polymerase sigma-70 factor family	81
3.4.4.	The <i>hasC</i> paralogs.....	82
3.4.4.1.	<i>S. equi</i> subsp. <i>equi</i> 4047 HasC paralogs.....	83
3.4.4.2.	<i>S. equi</i> subsp. <i>zooepidemicus</i> strain HasC paralogs.....	83
3.4.5.	The <i>murA</i> paralogs.....	84
3.4.6.	Conclusions	85
4.	Competing Metabolic Pathways of Hyaluronic Acid Synthesis	86
4.1.	Introduction.....	86
4.2.	Materials and Methods	89
4.2.1.	Preparation of antibiotic stocks.....	89
4.2.1.1.	Phosphomycin	89
4.2.1.2.	Ampicillin	89
4.2.2.	Antibiotic Minimum Inhibitory Concentration determination.....	89
4.2.3.	Capsule production analysis.....	90
4.2.4.	RNA extraction	90
4.2.5.	cDNA synthesis	92
4.2.6.	Primer design for RT PCR.....	92
4.3.	Results.....	93
4.3.1.	Phosphomycin MICs of <i>S. equi</i> strains of interest	93
4.3.2.	Capsule production in the presence of Phosphomycin	98
4.3.3.	RT-PCR analysis of <i>murA</i> paralog gene expression between phenotypes	99
4.4.	Discussion and Conclusions	99
4.4.1.	Phosphomycin MIC's in <i>S. equi</i> strains	100
4.4.2.	Capsule production in the presence of Phosphomycin	101
4.4.3.	Gene expression of the <i>MurA</i> paralogs between phenotypes	103
4.4.3.1.	Primer design	103
4.4.3.2.	PCR assay optimization	105
5.	Targeted proteomic comparison of <i>S. equi</i> strains	111
5.1.	Introduction.....	111
5.2.	Materials and Methods	113
5.2.1.	Protein extraction	113
5.2.2.	Proteomic analysis	114

5.3.	Results.....	115
5.3.1.	Whole cell protein profile of <i>S. equi</i> strains	115
5.3.2.	Protein identification in bands of interest.....	116
5.4.	Discussion	120
5.4.1.	Protein profiles of <i>S. equi</i> strains.....	121
5.4.2.	Protein identification in selected SDS-page bands.....	122
5.4.2.1.	Foldase PrsA.....	122
5.4.2.2.	Antiphagocytic M-like protein.....	123
5.4.2.3.	Superoxide dismutase.....	127
5.4.2.4.	General stress proteins.....	128
5.4.3.	Major surface structures of <i>S. equi</i>	129
5.4.3.1.	M-like protein and superoxide dismutase	129
5.4.3.2.	Foldase PrsA.....	130
5.4.4.	Conclusions	131
6.	Discussion and Conclusions.....	132
6.1.	Fermentation of <i>S. equi</i> species for HA production	132
6.1.1.	Advantages and disadvantages over animal production	132
6.1.2.	The fastidious nature of <i>Streptococcus</i>	133
6.1.3.	Current production procedure and its challenges	134
6.1.4.	<i>S. equi</i> subsp. <i>ruminantium</i> as a candidate for HA production.....	136
6.1.5.	Future work	136
6.2.	Comparative genomics for insight into HA synthesis regulation	136
6.2.1.	Challenges of DNA extraction of <i>S. equi</i> species.....	137
6.2.2.	Annotation, mis-annotation and assembly	138
6.2.3.	Selection of putative HA production-related targets	139
6.2.4.	Future work	139
6.3.	The HA production pathway in competition with essential processes	140
6.3.1.	The MurA enzymes as a target.....	141
6.3.2.	The use of phosphomycin for MurA inhibition	141
6.3.3.	Challenges due to paralog sequence similarity	142
6.3.4.	Future work	142
6.4.	Proteomic tools for distinguishing <i>S. equi</i> species	143
6.4.1.	Challenges of protein extraction of <i>S. equi</i> species	143
6.4.2.	Proteomics for subspecies differentiation	144
6.4.3.	Evidence of M-like protein processing or degradation.....	144
6.4.4.	Future work	145

6.5. Overall conclusions	145
Appendix A.	146
Appendix B.	147
Appendix C.	148
Appendix D.	149
References	150

Abbreviations

°C	degrees celsius
ANOVA	analysis of variance
AT	adenine and thymine
BHI	brain heart infusion
BLAST	basic local alignment search tool
bNAG	β - <i>N</i> -acetylglucosaminidase
CD44	cluster determinant 44
CD46	cluster of differentiation 46
CDM	chemically defined media
cDNA	complementary DNA
CSCs	cancer stem cells
Ct	cycle threshold
Cys115	cysteine at position 115
DNA	deoxyribonucleic acid
ECM	extracellular matrix
EDTA	ethylenediaminetetraacetic acid
ESI	electrospray ionisation
FM	fermentation media
G6P	glucose-6-phosphate
GAG	glycosaminoglycan
GAS	group A Streptococci
GC	guanine and cytosine
HA	hyaluronic acid
HARE	hyaluronan receptor for endocytosis

HAS	hyaluronan synthase
<i>hasA</i>	hyaluronan synthase
<i>hasB</i>	UDP-glucose dehydrogenase
<i>hasC/galU</i>	UDP-glucose pyrophosphorylase
<i>hasD/glmU</i>	N-acetylglucosamine-1-phosphate uridylyltransferase
<i>hasE/pgi</i>	phosphoglucoisomerase
HBA	horse blood agar
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HPLC	high performance liquid chromatography
kDa	kilodalton
LC-MS/MS	liquid chromatography-mass spectrometry
LPS	lipopolysaccharide
Lys22	lysine at position 22
LYVE-1	lymphatic vessel endothelial hyaluronan receptor
MDa	megadalton
Mga	M protein trans-acting positive regulator
MgCl ₂	magnesium chloride
MIC	minimum inhibitory concentration
Mn-SOD	manganese superoxide dismutase
MurA	UDP-N-acetylglucosamine enolpyruvyl transferase
NAD(P)	Nicotinamide adenine dinucleotide phosphate
NAG	N-acetylglucosamine
NaOH	sodium hydroxide
NCBI	national centre for biotechnology information
OD	optical density

OVD	ophthalmic medical device
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEP	phosphoenolpyruvate
PIPES	piperazine-N,N'-bis(2-ethanesulfonic acid)
PSA	polysialic acid
PTS	putative phosphotransferase system
RAST	rapid annotation using subsystem technology
RHAMM	receptor for hyaluronan-mediated motility
RNA	ribonucleic acid
RNS	reactive nitrogen species
ROS	reactive oxygen species
RPM	revolutions per minute
rRNA	ribosomal RNA
RT	reverse transcription
RT-PCR	reverse transcription polymerase chain reaction
SAXS	small-angle X-ray scattering
SDS	sodium dodecyl sulphate
SLA	sulphur limited agar
SLB	sulphur limited broth
SOCS3	suppressor of cytokine signaling 3
SOD	superoxide dismutase
TLR-4	toll-like receptor 4
T _m	melting temperature
UDP-GlcA	UDP-glucuronic acid

UDP-GlcNAc

UDP-*N*-acetylglucosamine

1. Introduction

1.1. Polysaccharides

Polysaccharides are generally large carbohydrate molecules composed of monosaccharide units and are found ubiquitously in living organisms, with some examples being guar gum, pectin and alginate (plants and algae), chondroitin and chitosan (animal) and xanthan gum and dextran (microbial) (Bhatia, 2016). They can be either linear in structure or possess branching groups of various size and structure and this variation in polysaccharide complexity correlates to the diverse functions observed such as energy storage, structural support, cellular protection and adherence to surfaces (Chakraborty *et al.*, 2019).

In prokaryotes, extracellular polysaccharides play a role in biofilm formation, cell migration and protection from environmental factors such as drying as well as protection from antimicrobial molecules and engulfment by phagocytes and predatory protozoa (Nwodo, Green and Okoh, 2012). Additionally, bacteria can produce polysaccharides as virulence factors such as lipopolysaccharide (LPS) in gram negative bacteria (Lukacova, Barak and Kazar, 2008), galactofuranose of *Mycoplasma* (Jores *et al.*, 2019), polysaccharide capsule of *Enterococcus faecium* (Ali, Blum and Sakiotanc, 2019) and polysialic acid (PSA) of group B streptococci, *N. meningitidis*, *Mannheimia haemolytica* and *E. coli* (Uchiyama *et al.*, 2019). Whilst the functions of extracellular polysaccharides are varied, some pathogenic bacteria have evolved to produce capsular polysaccharides which are identical to polysaccharides of their target host, such as *N. meningitidis* serogroup B, *H. influenza* (Hib) and some *Streptococcus* species. These capsules provide a mechanism for evasion of the immune system of the host (Cress *et al.*, 2014). Many of these bacterial polysaccharides possess properties which make them suitable for industrial and medical applications such as dextran, xanthan, cellulose and alginate which have applications in food thickeners, biomaterials in tissue engineering, wound dressings, chromatography and pharmaceuticals (Rehm, 2010). One such polysaccharide is hyaluronic acid which will be discussed further below and is produced by a variety of bacteria such as *Streptococcus pyogenes* (group A), *Streptococcus dysgalactiae* subsp. *equisimilis* (group B), subspecies of

Streptococcus equi (group C) and *Pasteurella multocida* (Necas *et al.*, 2008; Wajima *et al.*, 2016).

1.2. Hyaluronic Acid

1.2.1. Structure and Functions

Hyaluronic acid (HA) is a glycosaminoglycan (GAG) consisting of two repeating components D-glucuronic acid and *N*-acetylglucosamine connected respectively by β -1, 3 and β -1, 4 glycosidic bonds (Figure 1) (Liu *et al.*, 2011; Lu *et al.*, 2016). HA molecules are anionic in physiological pH (Tsepilov and Beloded, 2015). They are produced as high molecular weight chains which tangle together and result in highly viscous solutions which are non-Newtonian in nature (Cowman *et al.*, 2015). Non-Newtonian fluids are fluids which do not adhere to Newton's law of viscosity and as such the viscosity of these fluids may either increase or decrease when stress is applied (Denn, 2004). Solutions of HA demonstrate shear thinning, the loss of fluid viscosity when shear stress is applied such as passing the solution through a syringe, and at higher molecular weights the solutions also demonstrate elasticity resulting in HA being designated a pseudoplastic (Ambrosio *et al.*, 1999).

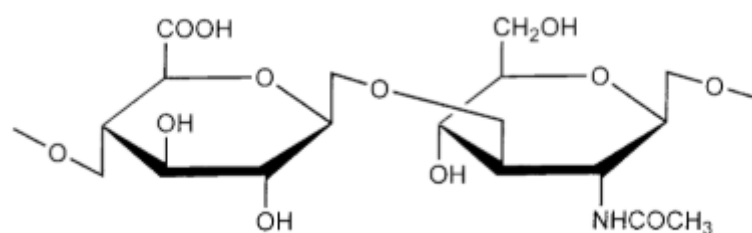


Figure 1 Hyaluronic acid disachharide. A section of hyaluronic acid chain comprised of precursors D-glucuronic acid (left) and *N*-acetylglucosamine (right) (Takehi, Kinoshita and Yasueda, 2003).

In both eukaryotes and bacteria HA is produced by membrane integrated glycosyltransferases called hyaluronan synthases (HAS) which assemble and extrude the chains in a progressive manner across the cell membrane. Whilst there are three isoforms of HAS in mammals they are structurally related to all other HAS enzymes bar the HAS of *Pasteurella multocida*.

Thus there are two classes of HAS enzymes with the *P. multocida* HAS being Class II and all other HAS enzymes being Class I (Weigel and DeAngelis, 2007). Within the body, HA is cleared from the ECM via the lymphatic system following either enzymatic degradation or fragmentation by reactive nitrogen and oxygen species. In an adult human, HA has a half-life of 2-6 minutes and the bulk of the elimination occurs within the liver, turning over between 10-100 mg of the total body concentration in a day (Necas *et al.*, 2008).

Whilst it is present in almost all vertebrate tissues the most significant quantities of HA are found within the extracellular matrix (ECM) and is most abundant within connective tissue, joints and skin where it functions as a molecular sieve and as a cushioning and lubricating agent. Its ability to function as these is considered to be due to its ability to bind up to 1,000 times its weight in water (Monslow, Govindaraju and Pure, 2015). High concentrations are also found within vitreous humour, umbilical cord and synovial fluid (Girish and Kemparaju, 2007). In fact it is one of the main components of synovial fluid (Tamer, 2013) and reduction in the elasticity and viscosity of the synovial fluid is attributed to the pathophysiology of osteoarthritis (Balazs, 2009). In addition to its role as a lubricant and space filler HA has been described as an organiser of the ECM and as a facilitator of cell migration through the provision of a suitable environment (Lee and Spicer, 2000). Certainly these functions support the observation that HA performs a role in the promotion of wound healing, which requires reorganisation of tissues and migration of immune cells thought to be facilitated by a sharp increase in HA concentration which is observed at the site of wounds (Aya and Stern, 2014). HA levels remain high in foetal wounds for a longer period of time than in adult wounds which is attributed to the lack of collagen accumulation and scar formation observed (Muto *et al.*, 2019). Wound healing is also facilitated by the binding of HA to molecules of the ECM and cell receptors, of which the major receptors are cluster determinant 44 (CD44), lymphatic vessel endothelial hyaluronan receptor (LYVE-1), hyaluronan receptor for endocytosis (HARE), receptor for hyaluronan-mediated motility (RHAMM) and toll-like receptor 4 (TLR4) (Solis *et al.*, 2012). CD44 is present on the surface of a number of

cell types such as epithelial and endothelial cells, fibroblasts and leukocytes and as such CD44 binding to HA is possibly the most well studied (Takehi, Kinoshita and Yasueda, 2003). Due to the presence of CD44 on leukocytes which are one of the main classes of immune cells, it was indicated that a possible role exists for HA in modulation of the immune system and it has since been identified that low and high molecular weight HA molecules produce different responses upon signalling through CD44, with low molecular weight HA fragments indicating injury and as such are pro-inflammatory and high molecular weight HA having the opposite effect (Marcellin, Steen and Nielsen, 2014). An example of this is stimulation of NF- κ B activity through CD44 signalling, which is promoted by HA fragments but inhibited by chains of $\geq 1 \times 10^6$ Da (Lee and Spicer, 2000). This is also the case for TLR4 where HA <900 kDa promote inflammatory responses whereas above this molecular weight the anti-inflammatory suppressor of cytokine signaling 3 (SOCS3) is upregulated (Asari, Kanemitsu and Kurihara, 2010).

Further to its role in cell signalling and migration, HA has been acknowledged to play a role in cancer development where it is speculated that the presence high concentrations of HA in the extracellular environment can contribute to facilitation of tumour angiogenesis and metastasis and as a result can effect prognosis of patients in a variety of cancer types (Chanmee, Ontong and Itano, 2016). Once again this function appears to be dependent on the molecular weight with it being speculated that HA molecules below 500 kDa promote the development of cancer stem cells (CSCs) which are associated with chemotherapy resistance (Price, Lokman and Ricciardelli, 2018), while higher molecular weight HA is thought to infer a resistance to cancer development, a theory thus far supported by the naked mole rat which is extensively studied for its resistance to cancer attributed to the high molecular weight HA found in its tissues (Tian *et al.*, 2013). RHAMM is a HA receptor speculated to be involved in tumour migration but with certain functions in tissue repair as it is present on the surface of endothelial cells (Jiang, Liang and Noble, 2011). Like with CD44, HA interaction with RHAMM can induce intracellular signalling which activates pathways involved in processes

such as cell trafficking and orientation, cell-cell interaction and growth and differentiation which all play roles in wound healing, inflammation and other pathologies (Misra *et al.*, 2015).

Thus HA provides a variety of functions in homeostasis and pathogenesis that are molecular weight dependent, making it a suitable candidate for investigation into treatment and prevention of a variety of conditions.

1.2.2. History and production methods

Hyaluronic acid was first identified and isolated from the vitreous humour of a cow in 1934 by Karl Meyer and John Palmer where it was named due to the presence of uronic acid in the molecule, its chemical structure was then resolved in the 1950's (Moscovici, 2015). It was first used commercially as an egg white substitute in 1942 prior to the knowledge of its structure, before being used clinically in the 1950's as an injectable substitution for the vitreous humour during eye surgery, a use it still holds to this day (Necas *et al.*, 2008).

Previous methods for production of HA have included extraction from human umbilical cords, bovine synovial fluid and vitreous humour and most recently extraction from rooster combs which contain a high concentration of HA, this technique is still being used (Kogan *et al.*, 2007). However an increasing awareness of the threat of cross species virus transmission and contamination with animal protein has prompted the investigation of other avenues, as such bacterial fermentation particularly with *Streptococcus* species, where HA is produced as a protective capsule, is the primary method of HA production (Boeriu *et al.*, 2013). Other avenues under investigation include the metabolic and genetic engineering of non-pathogenic strains such as *E. coli*, *B. subtilis* and *L. lactis* with the hyaluronic acid synthesis machinery from *Streptococcus* species. Some concern surrounding contamination from toxins in native HA producers prompted the development of these methods, however the molecular weight and concentrations are not yet adequate for production of HA for its many applications (Chong *et al.*, 2005). Due to the current production methods for HA for medical use, purification is of paramount importance in order to avoid contamination with pro-inflammatory molecules of either

animal or bacterial origin and often the purification processes are so rigorous that the ultimate yield and molecular weight of the product is affected (Choi *et al.*, 2014). Bacterial fermentation of HA will be discussed in more detail below.

1.2.3. Uses

As mentioned above HA is naturally present within the body, conserved between both eukaryotes and prokaryotes and possesses functions in a variety of biological processes involving cell migration and ECM remodelling. Due to this it is biocompatible and is used in a variety of ways which are discussed below.

1.2.3.1. Clinical

HA is used topically for wound healing with studies showing that application of HA to burn wounds shortens the healing time by promoting reepithelialisation through promotion of cell migration to the wound site (Aya and Stern, 2013). This promotion of healing is particularly useful in cases of impaired healing such as in aging, where senescence results in impaired cellular function (Fouda *et al.*, 2016) and diabetic wounds where healing is impaired due to a variety of factors including impaired collagen deposition, epidermal function, macrophage recruitment and angiogenic response (Brem and Tomic-Canic, 2007). HA produced by bacterial fermentation which was found to be free of problematic endotoxin concentration was found to accelerate wound contraction in rats in line with other claims of wound healing improvement upon topical application (Patil *et al.*, 2011).

As well as maintaining the viscosity and elastic properties of the synovial fluid within joints, HA binds and preserves aggrecan, a molecule which supports cartilage in its function of load bearing (Knudson *et al.*, 2019). Whilst the main cause of osteoarthritic symptoms is degradation of cartilage within the joint, the molecular weight of HA in the synovial fluid of arthritic joints is reduced to as low as 50% of that in healthy joints, a condition which is likely caused by degradation by reactive oxygen species (ROS) (Pontes-Quero *et*

al., 2019). Thus the injection of high molecular weight HA into the joints of patients has become an approach for relief of symptoms of osteoarthritis through replenishment of the viscoelastic properties of the synovial fluid and promotion of healing in a procedure termed viscosupplementation (Iannitti, Lodi and Palmieri, 2011). Often viscosupplementation involves the injection of HA in conjunction with corticosteroids in an attempt to alleviate pain and joint stiffness and to this day these therapies are under development to include aspects of cell therapy (Im, 2019).

As previously mentioned, the first clinical use of HA was as a cushioning agent for the vitreous humour during surgery. Whilst the concentration of HA in the eye varies greatly it is one of the main components of the vitreous humour and contributes to viscosity (Kleinberg *et al.*, 2011). While it can be used to treat eye wounds, the most common use of HA and its derivatives is in ophthalmic surgery such as cataract surgery and cornea replacement surgery, where it is applied in order to protect the delicate structures within the eye from damage and to regulate intraocular pressure (Kim *et al.*, 2017). The molecular weight of HA is an important factor in ophthalmic medical devices (OVD) as higher molecular weight increases the viscoelastic properties and so efforts to improve devices by enhancing the molecular weight of the HA used are ongoing (Balazs, 2009).

The fact HA is naturally occurring and thus is biocompatible and easily metabolised makes it suitable as a vehicle for drug delivery and due to its function as a ligand for CD44 receptors which are overexpressed on the surface of a variety of tumour cells it is particularly suited to targeted delivery for cancer therapies (Tripodo *et al.*, 2015). As such there has been a vast increase in literature over the last 25 years investigating the targeting capability of HA conjugated to either free drugs or carrier bound drugs (Dosio *et al.*, 2016) and the development of HA drug delivery vehicles paired with nanoparticles and other conjugates have been developed for delivery of chemotherapeutics such as doxorubicin as well as anti-viral drugs

like interferon α for targeted treatment of Hepatitis C (Lee *et al.*, 2012; Li *et al.*, 2016).

1.2.3.2. Tissue engineering

The cell signalling and ECM organisation capabilities of HA make it an attractive molecule for use as a biomaterial and as such HA scaffolds have become common place in tissue engineering (Gallo *et al.*, 2019). HA molecules can be cross linked and chemically modified to produce versatile hydrogels for use in a variety of fields including the modelling of biological tissues and environments (Trombino *et al.*, 2019). Such materials can be used to either replace entire or partial biological structures or aid in restoration of function with the more common cross linkages involving the carboxyl, hydroxyl and acetyl groups of the HA molecule (Hemshkhar *et al.*, 2016). HA hydrogels have also been used to model pathological tissues such as tumours in order to better assess drug efficacy (Arora *et al.*, 2017). Addition of HA to electrospun nanotube scaffolds designed for cartilage tissue engineering applications increases the hydrophilicity (Nikbakht *et al.*, 2019) and improved cell proliferation and regeneration of tissues has been observed when biological scaffolds are coated with HA (Tiwari, Patil and Bahadur, 2018). Thus the applications for HA in regenerative medicine and tissue engineering are numerous.

1.2.3.3. Moisturisers and cosmetics

Due to its hydrophilic nature, natural function as a space filler and free radical scavenging capability HA has been implemented in the cosmetics industry as a topical moisturiser in creams as well as components in fillers (Fallacara *et al.*, 2018). HA was found to be more effective than collagen at preserving cosmetic corrections when used in facial fillers (Kogan *et al.*, 2007) and HA based fillers scored higher in client satisfaction than others as the flexibility of HA fillers produces results which are more natural in appearance (Solish *et al.*, 2019). The use of HA subdermal fillers results in effects lasting approximately 6 months and is currently the most common type of

filler, whilst increasingly also being used in moisturisers and skin treatments (Salwowska *et al.*, 2016).

1.3. Microbial Fermentation for HA production

Microbial fermentation is now the most common method for production of HA, most often using species of Group C streptococci (Chen *et al.*, 2012). Due to the rheological properties of HA it is currently only feasible to produce maximum concentrations of 5-10 g/L as the broth viscosity is too high for successful mass transfer above this, leading to inadequate dissolved oxygen concentration (Blank, McLaughlin and Nielsen, 2005). HA fermentation is carried out in bioreactors under varying culture conditions which can influence the concentration and molecular weight of the product, such as temperature, pH level, dissolved oxygen content, metal ion content and nutrient types and availability (Pires, Eguchi and Santana, 2010). Group C streptococci are fastidious bacteria and demanding with regards to nutrient requirements for growth (Armstrong, Cooney and Johns, 1997). The most commonly used fermentation conditions for HA production to date include a temperature range of 30-37°C, a pH range of 6.5- 7.5 and a carbon source of glucose usually at concentration of 10-60 g/L (Armstrong and Johns, 1997). Production of HA in these bacteria competes with growth and so fermentation conditions are often set to provide a tolerable stress, such as inadequate nutrient supply, in order to optimise the HA production without significantly inhibiting bacterial growth (Boeriu *et al.*, 2013). Group C streptococci are lactic acid producing bacteria and during fermentation approximately 80-85% of the carbon supply is utilised for lactic acid and acetic acid production with only 5-10% reserved for growth and capsule production (Lu *et al.*, 2016). In addition, increasing lactic acid concentrations inhibit growth and HA production and as such are a major hindrance in the fermentation process for HA production (Duan *et al.*, 2008). Capsule synthesis and lactic acid production both occur during log phase of fermentation, whereas degradation of HA often occurs during stationary phase (Wu *et al.*, 2009).

The most common mode of fermentation is batch or fed batch, where the bacteria are fermented within a closed system until stationary phase is reached and the product is harvested. This method is laborious due to the need for regular bioreactor turnover and it is speculated that continuous culture would

prevent stationary phase from being reached thus increasing yield and reducing the risk of degradation of the product (Chong *et al.*, 2005). However in practice a decline in HA production is observed independent of continuation of growth suggesting instability of HA production in continuous culture where a continuous supply of nutrients is provided to prevent stationary phase being reached. As a result, batch fermentation remains the primary mode for HA production (Liu *et al.*, 2011). Efforts to produce hyaluronic acid by microbial fermentation using cheaper and more environmentally conscious methods have been made using soy, derivatives from agricultural sources, by products and discards from the fishing industry and even cashew apple juice (Pires *et al.*, 2010; Benedini and Santana, 2013; Oliveira *et al.*, 2013; Vazquez *et al.*, 2015). More recently a two-stage fermentation process was developed using a strain of *S. equi* subsp. *zooepidemicus* in order to optimise both molecular weight and yield as previously studies have focused on one or the other (Liu *et al.*, 2018). Due to the pathogenic nature of Streptococci which will be discussed below there are some concerns around the use of these bacteria for HA production and so efforts to genetically engineer safer alternatives are ongoing.

L. lactis is a species of bacteria which is commonly engineered to produce HA, however it was observed that the production of lactic acid resulted in reduced HA yield and molecular weight (Chauhan *et al.*, 2014). This challenge was overcome by engineering a lactate dehydrogenase-deficient strain which resulted in an increased production of ethanol and acetoin as well as a three-fold increase in both the molecular weight and yield of HA (Kaur and Jayaraman, 2016). However, although substantially increased in comparison to other *L. lactis* strains both the yield and molecular weight produced remain lower than that achieved by natural HA producing species. The same has been observed for *E. coli* strains engineered to possess the genes involved in HA production from *S. pyogenes* (Yu and Stephanopoulos, 2008). The hyaluronan synthase gene from *S. equi* subsp. *zooepidemicus* was cloned into *B. subtilis* and rapidly mutated in order to produce a number of variants for screening for increased HA yield and molecular weight (Zhang *et al.*, 2016), however once again the highest yield and molecular weight achieved were substantially lower than that achievable with Streptococci.

Therefore the replacement of *Streptococcus* species as the primary microbe for HA fermentation still requires vast improvements in order to be feasible.

1.4. Streptococci

Streptococci are a genus of gram-positive bacteria which are spherical in shape and grow in long chains. The majority of the species are facultative anaerobes, do not produce catalase and have stringent and varying requirements with regards to nutrition and environmental conditions (Timoney, 2010). This genus is classed as one of the most invasive in terms of disease, with 35 out of the 49 species having been identified as causing invasive infection in humans, the most common of which are *S. pyogenes*, *S. agalactiae* and *S. mutans* (Krzyściak *et al.*, 2013). *S. pyogenes* is a member of Group A streptococci which is a common cause of disease and mortality throughout the world, responsible for a variety of conditions from pharyngitis to toxic shock syndrome, necrotizing fasciitis and sepsis (Lamagni *et al.*, 2008). Conditions associated with this bacteria such as Scarlet Fever are on the increase, particularly in children of school age, who are the most common asymptomatic carriers. Penicillin is the main method of treatment, however reports of resistance are now common (Wong and Yuen, 2012). *S. pyogenes* is one of the species which produce HA as a protective capsule (Kang *et al.*, 2012) however due to its pathogenicity it is not suitable for use in industrial production of HA.

In addition to Group A streptococci, HA is produced by others, notably *S. uberis* (a major cause of mastitis in dairy cows) (Field *et al.*, 2003) and several Group C Streptococci which will be discussed below.

1.4.1. Group C Streptococci

Group C Streptococci are a group of gram-positive, facultative anaerobic, catalase negative and pathogenic bacteria which were previously described as comprising of four species, *S. dysgalactiae* subsp. *dysgalactiae*, *S. constellatus* subsp. *pharynges*, *S. equi* subsp. *equi* and *S. equi* subsp. *zooepidemicus* (Klos and Wojkowska-Mach, 2017). These bacteria are capable of causing disease in both humans and numerous animal species such as horses, pigs, ruminants, dogs, birds, dolphins and rodents with some subspecies showing host-specificity. They have also been identified amongst normal mucosal flora of all types of host (Sinner and Tunkel,

2015). Infection is associated with the presence of virulence factors such as M-like protein, capsule, streptokinases, pyrogenic exotoxins, haemolysin and other proteases (Segura and Gottschalk, 2005).

Subspecies of *Streptococcus equi* are responsible for severe instances of disease in domesticated animals. *S. equi* subspecies *equi* is the known causative agent of a frequently occurring upper respiratory disease in horses named Strangles. Strangles is highly infectious, varies in severity and in a number of cases infection persists following resolution of symptoms, often due to lack of clearance in the guttural pouch (Timoney, 2004; Waller, Paillot and Timoney, 2011). The high risk of complications and lengthy quarantine of infected yards during an outbreak of the disease causes significant economic loss. *S. equi* subsp. *equi* is restricted to the horse as its host, however is not the only member of *S. equi* to cause disease in horses. *S. equi* subspecies *zooepidemicus* is a commensal organism found within the natural flora of the horse. Mucosal disease is associated with opportunistic infection, and can also occur in other animals such as cattle, sheep, pigs, cats, dogs and goats (Steward *et al.*, 2017).

Instances of serious and invasive disease in humans have been reported in connection with this bacteria to such an extent that it was proposed to be an emerging problematic zoonotic (Pelkonen *et al.*, 2013). Certainly there is evidence to suggest that transmission to humans may occur from horses, dogs and ruminants either by direct contact or through the consumption of unpasteurised dairy products (Abbott *et al.*, 2010). This is not an implausible concept when considering the extent of similarities to the previously mentioned human pathogen *S. pyogenes*. They share a number of genes and traits considered to pertain to virulence such as those for surface M proteins, antigens and the HA polysaccharide capsule. This trait similarity is speculated to be due to their origination from a common ancestor (Klos and Wojkowska-Mach, 2017). Additionally, mastitis in cows caused by *S. pyogenes* has been documented and thought to have arisen by transmission from humans to the animal (Wong and Yuen, 2012). *S. equi* subsp. *equi* is proposed to be descended from an ancestral strain of *S. equi* subsp. *zooepidemicus* and the two species share numerous virulence factors such as the ones listed above (Timoney, 2004). The suggestion that

these two species are descended from a common ancestral strain would certainly explain the >98% genomic identity that is observed between them. In addition to this, *S. equi* subsp. *zooepidemicus* shares >80% of its genetic identity with *S. pyogenes* (Holden *et al.*, 2009), a trait which correlates with the observation that both can cause severe infections in humans such as rheumatic fever and glomerulonephritis (Pelkonen *et al.*, 2013).

In 2004, previously unclassified gram positive cocci with >70% genetic identity with *S. equi* subsp. *equi* and *S. equi* subsp. *zooepidemicus* was isolated during a routine bacteriological study investigating causative agents of mastitis in small ruminants. Following phenotypic, biochemical and genetic analysis the organism was proposed to be a new species and designated *Streptococcus equi* subsp. *ruminatorum* (Fernandez *et al.*, 2004). Following this publication the strain was attributed to Strangles like disease affecting spotted hyenas and plains zebras in Tanzania, where several clones of the strain were identified throughout the wild population (Speck *et al.*, 2008). That same study attempted to establish differences and similarities of these wildlife strains to the originally isolated strain by Fernandez *et al.* (2004) as well as type strains of *S. equi* subsp. *equi* and *S. equi* subsp. *zooepidemicus* using biochemical and molecular biological methods. Differences and similarities were reported however unfortunately not all data was presented. Little more has been published about this novel subspecies of *S. equi*.

1.4.2. Hyaluronic Acid Capsule of Streptococci

1.4.2.1. Hyaluronic acid protection against phagocytosis

For all of the *Streptococcus* species discussed above, the hyaluronic acid capsule is considered to be one of the main factors affecting virulence and pathogenicity (Wessels *et al.*, 1991; Anzai *et al.*, 1999; Dinkla *et al.*, 2007). It is speculated that production of capsule enables the bacteria to evade host immune responses such as identification by immune cells and deposition of complement factors allowing them to more effectively establish infection (Cress *et al.*, 2014; Fischetti and Dale, 2016). Due to the natural presence of HA in

the ECM of the host, capsule produced by Streptococci behaves as a form of invisibility cloak, preventing identification from immune cells which will view the HA as 'self' and the presence of HA correlates positively with a reduced number of bacteria cells which undergo phagocytosis (Schommer *et al.*, 2014).

As previously mentioned, HA is one of the main virulence factors present in *S. equi* species, along with haemolytic exotoxin streptolysin S and the antiphagocytic M-like protein (Timoney, 2004).

Antiphagocytic M-like protein functions by impeding the deposition of opsonic C3b molecules to the bacterial surface, resulting in reduced formation of membrane attack complex and thus reduced complement mediated bacterial death (Merant, Sheoran and Timoney, 2011) and both it and HA are required for resistance to phagocytosis (Dale *et al.*, 1996a). While it is speculated that HA protects against phagocytosis by physically hindering opsonins and cell binding proteins as a result of its polyanionic and hydrophilic qualities which prevent penetration through the HA chains surrounding the bacteria cells (Dinkla *et al.*, 2007), this is not the generally accepted mechanism and there is the opinion that HA acts to physically inhibit the contact of phagocytes with deposited opsonins (Nizet, 2007). Additionally it has been observed that encapsulated bacteria form aggregates which may aid in protection from reactive oxygen species produced during immune responses by reducing their infiltration (Henningham *et al.*, 2015).

1.4.2.2. Genetic basis of capsule synthesis

Among Group C Streptococci, *S. equi* subsp. *zooepidemicus* has served as prototype for HA biosynthesis and regulation. The production of HA involve five genes, the hyaluronan synthase (*hasA*) which is responsible for the assembly of the HA chain, UDP-glucose dehydrogenase (*hasB*) and UDP-glucose pyrophosphorylase (*hasC/galU*) which are responsible for the synthesis of the precursor UDP-glucuronic acid and *N*-acetylglucosamine-1-phosphate uridylyltransferase (*hasD/glmU*) and phosphoglucoisomerase (*hasE/pgi*) which are responsible for the synthesis of UDP-*N*-acetylglucosamine (Chen *et al.*, 2014). The *has* operon of *S. equi*

subsp. *equi* is interrupted by an insertion of approx. 11200 base pairs encoding a putative phosphotransferase system (PTS) between the first 3 genes (*hasA*, *hasB* and *hasC*) and *hasD* and *hasE* which are present within a separate operon, whereas in *S. equi* subsp.

zooepidemicus all five genes are located within a single *has* operon.

In addition in, a second copy of *hasC* is found in a separate operon in both strains along with a gene for NAD(P)⁺ dependent glycerol-3-phosphate dehydrogenase (Blank, Hugenholtz and Nielsen, 2008).



Figure 2 *has* operon of *S. equi* species. The *has* operon of *S. equi* subsp. *zooepidemicus* contains all five genes involved in HA synthesis. In *S. equi* subsp. *equi* the operon has been interrupted, separating *glmU* (*hasD*) and *pgi* (*hasE*) from *hasA*, *hasB* and *hasC* (Blank, Hugenholtz and Nielsen, 2008).

A study assessing the effects of over expressing the 5 genes of the *has* operon using a nisin inducible vector in a *S. equi* subsp.

zooepidemicus strain on the yield and molecular weight of HA

capsule revealed the importance of individual genes (Chen *et al.*,

2009b). Overexpression of the *hasA* gene increased the yield of HA,

however reduced the molecular weight in comparison with the

wildtype strain. Contrastingly overexpression of *hasC* resulted in a

decreased yield as did expression of both *hasD* and *hasE*

simultaneously. Additional significant findings of this study were that

overexpression of *hasE* increased the molecular weight of the HA

product as did the presence of the empty vector plasmid control.

The above outcomes prompted an investigation into the effects of the

empty plasmid on gene expression and hyaluronic acid synthesis

(Marcellin, Chen and Nielsen, 2010). The unexpected result of this

follow on study was that a downregulation of MurA activity was

identified and further investigation concluded that overexpression of

this enzyme caused the molecular weight of HA produced to decrease.

MurA is an enzyme which catalyses the first step in the pathway of

peptidoglycan synthesis. This step involves the relocation of an

enolpyruvate residue from phosphoenolpyruvate (PEP) to position 3

of UDP-*N*-acetylglucosamine (Zoeiby, Sanschagrin and Levesque, 2003). As mentioned previously, *N*-acetylglucosamine is one of the components of the hyaluronic acid chain and thus this step of the peptidoglycan synthesis pathway is in direct competition with hyaluronic chain synthesis. Thus it brings to reason that down-regulation of the MurA enzyme would enhance availability of the *N*-acetylglucosamine supply for hyaluronic acid synthesis. Prior to the identification of this effect of MurA downregulation it had been concluded that HA molecular weight was dependent on the ratio of available precursors and that *N*-acetylglucosamine was the limiting component (Chen *et al.*, 2009b).

1.4.2.3. Hyaluronan synthase

As mentioned above the hyaluronan synthase enzyme is responsible for the synthesis of the HA chain. This enzyme is a membrane-integrated glycosyltransferase which extrudes the HA chain through the bacterial membrane as it assembles it (Weigel, 2015). Only hyaluronan synthase is required in order to assemble the HA chain and this is due to the fact the HAS enzyme functions with two different glycosyl transferase activities (Weigel, 2002). The structure and topology of the HAS enzyme within the membrane has been elucidated (Heldermon, DeAngelis and Weigel, 2001) and mutagenesis studies on the *hasA* enzyme have been attempted in various strains in order to manipulate the yield and molecular weight of HA (Heldermon *et al.*, 2001; Weigel and Baggenstoss, 2012). In addition it has been observed that the enzyme is associated with molecules of cardiolipin and that this interaction is required for synthase activity (Tlapak-Simmons *et al.*, 1998). Cardiolipin is a phospholipid found in the membranes of mitochondria of eukaryote cells and bacteria which has been associated with the stability of numerous other integral membrane proteins (Musatov and Sedlák, 2017). However further investigations into the importance of this relationship are yet to be published.

1.5. Project aims

With HA being an increasingly important biomedical product as well as a virulence factor for this pathogenic species of bacteria, its role in the bacterial metabolism, infection establishment and the factors which influence its production are key topics for research. Previous investigations into HA production in *Streptococcus* focus on fermentation parameters and mutation of the genes within the *has* operon and often focus on a single strain. This study aims to investigate HA production from the perspective of the evolutionary necessity for streptococcal synthesis of HA and to use mucoid and non-mucoid strains to attempt to explore factors out with the immediate HA synthesis machinery. Specific aspects addressed are:

- Evaluation of current procedures for HA production from *S. equi* subsp. *equi*.
- Comparison of HA production in alternative Group C *Streptococcus* subspecies.
- Genomic comparison of *S. equi* strains and mucoid/non-mucoid phenotypes.
- Assessment of competing essential metabolic processes on HA production.
- Transcriptional and proteomic variations between Group C *Streptococcus* subspecies and HA producing and non-producing variants.

2. Growth Parameters and Permutations in HA production by *Streptococcus equi* subspecies

2.1. Introduction

Although it is more efficient and ethical than extraction from animal tissues, bacterial production of HA poses some challenges, including culture of source organisms. Lancefield group C bacteria such as *S. equi* subsp. *equi* and *S. equi* subsp. *zooepidemicus* are the most common native producing bacteria used to produce HA (Kim *et al.*, 1996; Chen *et al.*, 2009a; Choi *et al.*, 2014). While batch culture appears to be the most conventional mode of fermentation for HA production, there are arguments that continuous culture prevents progression into stationary phase thus avoiding the production of degrading enzymes and release of contaminating intracellular proteins and toxins which make the purification process more complicated (Chong *et al.*, 2005). However, in practice continuous culture has been unattainable due to the instability of the HA producing phenotype and subsequent tendency for HA production to decline over time, therefore the prevailing method of fermentation for HA production remains batch culture (Armstrong and Johns, 1997; Liu *et al.*, 2011). A stable continuous culture of *S. equi* subsp. *zooepidemicus* ATCC 35246 was developed which operated at a dilution rate of 0.4 h⁻¹ in an arginine limited chemically defined media (CDM) for 74 hours before a non-mucoid phenotype emerged. However CDM is regarded as too expensive for use in commercial production (Blank, McLaughlin and Nielsen, 2005). Fermentation conditions including temperature, pH and oxygen and nutrient availability can all have an impact on product characteristics and yields and they need to be monitored and controlled (Pires, Eguchi and Santana, 2010).

Continuous culture is the method currently employed by Hyaltech Ltd. This company uses NCIMB *S. equi* strain 40327 (designated SE40327 in this study) for production of hyaluronic acid to be used in medical solutions for ophthalmic viscosurgical devices (OVDs) for cataract surgery, bone fillers and viscosupplements used in osteoarthritis treatment. Bacterial fermentation is performed in volumes of between 80-130 L of the fermentation media described in the methods section (Table 1). Use of the strain for HA production was patented in 1992 (International Publication Number WO1992008799A1). As a result no available data exists that could be used to compare it to other strains of

group C Streptococci in regards to HA production, therefore it was decided that this study would focus on characterisation of SE40327 as a HA producer in comparison to other encapsulated strains.

The following strains were acquired from DSMZ for comparison to SE40327. DSMZ-20561 is a type strain of *S. equi* subsp *equi* which was isolated from a submaxillary abscess of a foal displaying symptoms of Strangles. DSMZ-20727 is a type strain of *S. equi* subsp *zooepidemicus* isolated from a case of bovine mastitis. Both of these strains are recommended by ATCC for use as quality control strains. DSMZ-17037 was first isolated in 2004 from the mastitis of ruminants and following phenotypic and molecular analysis was proposed to be a novel subspecies of *S. equi*, designated *S. equi* subsp. *ruminatorum* (Fernandez *et al.*, 2004). All three type strains were biochemically analysed and found to produce capsule (Speck *et al.*, 2008) and so were considered suitable candidates for comparison.

It is argued that continuous culture is the most suitable way of studying the physiology of bacteria without the interference from external growth determinants such as the eventual limitation of nutrients which induce the development of stationary phase observed in batch culture (Hoskisson and Hobbs, 2005). However, it has been claimed that this mode of culture is not a reliable approach for HA production due to the previously mentioned tendency of naturally HA producing bacteria to switch to a non-mucoid phenotype. This is a drawback that is seen in practice with SE40327, with the concentration of HA declining over time regardless of the stability of the biomass in a continuous system. An outcome which suggests that not all changes to growth physiology can be prevented by the continuous culture mode.

Whilst continuous culture and the use of chemostats is most certainly an important tool for assessing the effects of culture conditions on growth and production of molecules by bacteria, even in smaller vessels they can require large volumes of media and can be time consuming to set up and monitor. For the purpose of this work a small scale model of batch fermentation was developed using multi-well plates to compare strains of group C *Streptococci* in regards to HA production, as even shake flask culture can swiftly use up media resources. By comparison, use of multiwell plates (e.g. 96 well plates)

enhances experimental multiplexing, replicability and sample monitoring across multiple parameter permutations.

The fermentation parameters which influence HA production have been the focus of many research studies already (Kim *et al.*, 1996; Armstrong, Cooney and Johns, 1997; Armstrong and Johns, 1997; Blank, McLaughlin and Nielsen, 2005; Kim, Park and Kim, 2006; Duan *et al.*, 2008; Huang, Chen and Chen, 2008; Liu *et al.*, 2008; Chen *et al.*, 2009a; Zhang, Duan and Tan, 2010; Chen *et al.*, 2012). However, despite use in biomanufacturing, no data or protocols were readily available for SE40327 growth and HA production in batch mode or in small scale culture.

For fermentation of SE40327 the culture temperature and pH used are 37°C and 6.1. This pH was considered a low starting pH for batch culture as group C streptococci are lactic acid producing bacteria (Smith and Sherman, 1942; Papadimitriou *et al.*, 2016) which were expected to rapidly reduce the pH of the culture media. So experiments were proceeded at pH 7 which is commonly used for HA production in other strains of Streptococci (Liu *et al.*, 2011).

The set of experiments described in this chapter aimed to develop a reproducible and reliable protocol that would allow comparisons of HA producing strains in a small scale 96 well format. This adapted the fermentation media normally used in continuous fermentation of SE40327, optimising it for use in small scale batch conditions and comparing it to other media types to find the optimal conditions for growth and HA production for each subspecies.

2.2. Materials and Methods

2.2.1. Culture media and Bacteria Strains

2.2.1.1. Fermentation media

Fermentation Media (FM) and Sulphur Limited Broth (SLB) were prepared in accordance with standard operating procedures acquired from Hyaltech Ltd containing the concentrations below. All ingredients were provided by Hyaltech Ltd unless otherwise stated as were solutions of trace salts, sodium molybdate and calcium carbonate. Original supplier details for these ingredients are not

available. Hydrochloric acid was sourced from Merck. The ingredients were dissolved one at a time in the order below in deionised water with stirring. Fermentation Media was filter sterilised using Millipore 0.22 µm Steritop and Stericup. Sulphur Limited Broth was autoclaved at 121°C for 15 minutes. The pH of each was adjusted prior to sterilisation by addition of 16% NaOH. Fermentation Media was adjusted to pH 7 and Sulphur Limited Broth to pH 6.6.

Table 1. Fermentation media

Ingredient	g/L
Citric Acid	0.5
Sodium Dihydrogen Phosphate	2.4
Ammonium Chloride	2.5
Potassium Chloride	0.8
Magnesium Oxide	0.5
Hydrochloric Acid	2.4
Calcium Carbonate Stock	1.2
Sodium Molybdate Stock	1.1
Trace Salts Stock	0.5
Yeast Extract	7.5
Glucose	74

Table 2. Sulphur Limited Broth

Ingredient	g/L
Hydrochloric acid	3.9
Magnesium Oxide	0.7
Ammonium Chloride	75
Citric Acid	6
Sodium Dihydrogen Phosphate	40
Potassium Chloride	5
Potassium Dihydrogen Phosphate	19
Dipotassium Hydrogen Phosphate	22

Ingredient	mL/L
Sodium Molybdate Stock	1.5
Calcium Carbonate Stock	15
Trace Salts Stock	7.5

Table 3. Trace Salts Stock Solution

Ingredient	g/L
Zinc Oxide	1.6
Ferric Chloride Solution	45.9
Manganous Chloride	8
Cupric Chloride	0.7
Cobaltous Chloride	1.9
Hydrochloric Acid	17.2

Table 4. Calcium Carbonate Stock Solution

Ingredient	g/L
Calcium Carbonate	10
Hydrochloric Acid	23.2

Table 5. Sodium Molybdate Stock Solution

Ingredient	g/L
Sodium Molybdate	4.8
Hydrochloric acid	As required

Table 6. Sulphur Limited Broth

Ingredient	g/L
Stock Solution for SLB	70
Yeast Extract	5
Glucose	2.4

2.2.1.2. Brain Heart Infusion Media and Agar plates

Brain Heart Infusion (BHI) media and agar were sourced from Sigma-Aldrich Ltd and prepared in accordance with the instructions with one amendment. The BHI media was filter sterilised using Millipore 0.22 µm Steritop and Stericup (Merck). Sulphur Limited Agar (SLA) and 5% Horse Blood Agar plates were sourced from Southern Group Laboratory Ltd.

2.2.1.3. Bacteria Strains

S. equi NCIMB 40327 (SE40327) was received from Hyaltech Ltd as colonies on sulphur limited agar plates. *S. equi* subsp. *equi* DSMZ 20561 (SEE20561), *S. equi* subsp. *zooepidemicus* DSMZ 20727 (SEZ20727) and *S. equi* subsp. *ruminantium* DSMZ 17037 (SER17037) were acquired from DSMZ as ampoules of freeze dried culture which were revived in BHI media overnight at 37°C. Frozen 25% glycerol stocks of all strains were prepared by mixing 1ml of 50% glycerol solution with 1ml of bacteria culture in BHI media. Glycerol stocks were stored frozen at -80°C.

2.2.2. Establishing growth in fermentation media in comparison to other high nutrient media

2.2.2.1. Revival of Glycerol Stocks and Inoculum preparation

Frozen glycerol stocks of bacteria were removed from -80 freezer and streaked onto sulphur limited agar plates. The plates were incubated at 37°C for 48-72 hours in accordance with Hyaltech standard operating procedure. Single mucoid colonies were selected and inoculated into 50ml pre-equilibrated, sterile sulphur limited broth. The broth cultures were incubated at 37°C overnight prior to use.

2.2.2.2. Multi-well plate growth curve set up

Inocula were prepared as described previously (2.2.2.1) for the four strains of *S. equi* described above. All media was equilibrated to 37°C prior to use. The optical density (OD) at 660 nm was measured for each overnight culture and was diluted to 0.05 in the selected media. The samples were then seeded into a flat bottomed, untreated 96-well

plate in replicates of 15. Each well was filled with 200 μ l of sample and uninoculated media was used for blank wells. Plates were covered with a Breathe-Easy® sealing membrane (Sigma-Aldrich) and placed in a FLUOstar Omega Microplate reader (BMG Labtech). The plate was incubated at 37°C for 20 hours. Optical density readings at 660 nm were taken every 30 minutes with 200 RPM of shaking for 1 second prior to each reading. Data was plotted as averages and standard deviations of blank corrected time point OD readings.

2.2.3. pH control using HEPES, PIPES and Tris buffers

Inocula were prepared as described previously (2.2.2.1) with *S. equi* NCIMB 40327. The OD of the inoculum was measured at 660 nm and corrected to 0.05 in fermentation media supplemented with concentrations of 0, 10, 25 and 50 mM HEPES buffer solution (1M, pH 7-7.2) (Sigma-Aldrich). Each sample was seeded (200 μ l in each well) in replicates of 15 and uninoculated, buffered media was used for blank wells. Plates were incubated in a FLUOstar Omega Microplate reader (BMG Labtech) under conditions described previously (2.2.2.2). Data were plotted as averages and standard deviations of blank corrected time point OD readings. The same protocol was used for PIPES (1M, pH 6.0) (VWR) and Trizma base (1M in H₂O, pH 10.5-12) (Sigma-Aldrich). Dilutions were carried out as follows. To obtain media with 10, 25 and 50 mM buffer concentrations a 1M stock solution was diluted 1/10, 1/4 and 1/2 in media to produce concentrations of 100, 250 and 500 mM. These were then diluted 1/10 to obtain media for inocula at concentrations of 10, 25 and 50 mM.

2.2.4. pH monitoring of *S. equi* cultures

Inocula were prepared as described previously (2.2.2.1) with *S. equi* NCIMB 40327. A 1 in 10 dilution of the inoculum in 37°C fermentation media (pH 7) was prepared in a series of tubes, 3 for each 2 hour time point. The tubes were incubated at 37°C at 150 RPM shaking, the lids were left loose to allow for some gas transfer. Every 2 hours the optical density was measured by removing a set of tubes and transferring media to cuvettes. The pH of the culture in each tube was then measured using a pH

meter. The cultures were discarded following pH measurement. Growth and pH were monitored for 8 hours, where stationary phase had previously been shown to have been reached.

2.2.5. Capsule production analysis

Inocula for all four *Streptococcus equi* strains were prepared as described previously (2.2.2.1) and multi-well plates set up for growth in fermentation media (pH 7) as described previously (2.2.2.2). Optical density readings at 660nm were carried out against a blank of un-inoculated sulphur limited broth. The optical density for each culture was adjusted to 0.05 in fresh medium as described previously (2.2.2.2). The cultures were seeded 200 μ L per well in a 96 well plate along with single well blanks of 200 μ L un-inoculated media per strain. Plates were then sealed with a Breathe-Easy® sealing membrane (Sigma-Aldrich), placed in either the Omega FLUOstar plate reader and incubated until stationary phase was reached (max 8 hours) under conditions described previously (2.2.2.2). Following 8 hours of growth the plate was removed and the cultures pooled. The cultures were drawn into a 5 mL syringe and filtered through 1.0 μ m and 0.45 μ m syringe filters (Whatman®) consecutively. The supernatant samples were stored at -20°C until thawed for analysis by HPLC using a size exclusion column which was carried out at Hyaltech Ltd as per their standard operating procedure where retention time and HA concentration are measured for quality control purposes.

2.3. Results

2.3.1. Colony morphology of SE40327 compared to other *S. equi* strains

The colony morphology of SE40327 was compared to the other *S. equi* strains on sulphur limited agar plates and 5% horse blood agar plates. On sulphur limited agar all strains grew as translucent to opaque colonies, however while SE40327 and SER17037 colonies were large, rounded and glossy in appearance both SEE20561 and SEZ20727 appeared matte and less domed in shape (Figure 3). On horse blood agar the colonies of all strains appeared glossy in appearance and all displayed β -haemolysis. The strain which most closely matched SE40327 in morphology in both cases was SER17037.

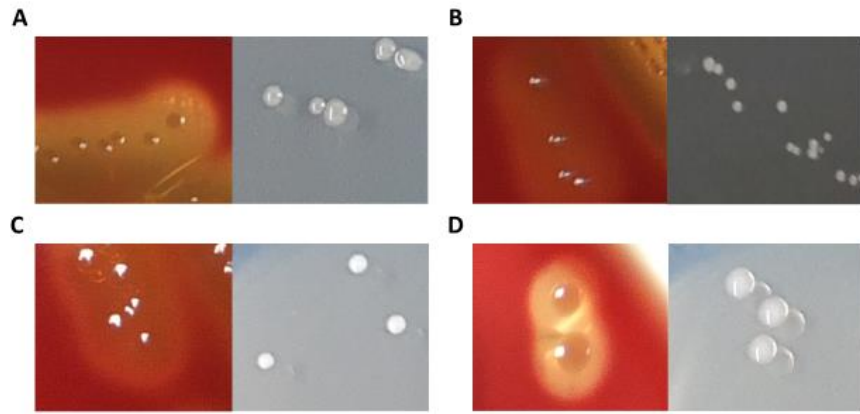


Figure 3 Colony morphology of the four selected strains of *S. equi* on 5% horse blood agar plates (left) and sulphur limited agar plates (right). The strains pictured are SE40327 (A), SEE20561 (B), SEZ20727 (C) and SER17037 (D). Bacterial stocks were streaked onto agar plates which were incubated for 48 hours at 37 °C.

To assess whether the use of sulphur limited agar promotes capsule production, the strains were grown on both SLA plates and BHI agar plates in order to compare the colony morphology. Figure 4 shows SE40327 cultured on both media. From the pictures it is apparent that whilst the colonies on SLA are indeed mucoid in appearance, they are not as large or glossy as on BHI media. Growth on BHI agar was also much more rapid than SLA, with distinct colonies being visible after 24 hours and colonies appearing to begin drying after 48 hours. The same observation can be made for the other 3 strains (figures 5-7), with colonies appearing larger and mucoid on BHI compared to SLA after 48 hours of growth. Of particular interest is SER17037 in Figure 5 which appears to display both mucoid and non-mucoid variant colonies on both SLA and BHI agar. Non-mucoid colonies appear much smaller and matte white in colour. The mucoid colonies on SLA are larger and appear domed and glossy, whilst on

BHI agar single colonies are indiscernible with growth appearing as areas of slimy, mucoid puddles.



Figure 4 Colony morphology of *S. equi* NCIMB40327 on Sulphur limited agar (left) and Brain Heart Infusion agar (right). Bacterial stocks were streaked onto agar plates and incubated for 48 hours at 37°C.

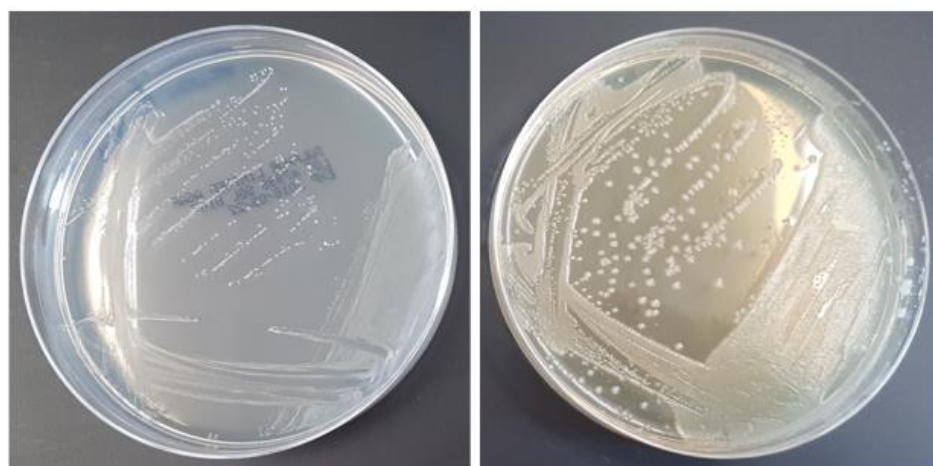


Figure 5 Colony morphology of *S. equi* subsp *equi* 20561 on sulphur limited agar (left) and brain heart infusion agar (right). Bacterial stocks were streaked onto agar plates and incubated for 48 hours at 37°C.



Figure 6 Colony morphology of *S. equi* subsp *zooepidemicus* 20727 on sulphur limited agar (left) and brain heart infusion agar (right). Bacterial stocks were streaked onto agar plates and incubated for 48 hours at 37 °.



Figure 7 Colony morphology of *S. equi* subsp *ruminatorum* 17037 on sulphur limited agar (left) and brain heart in fusion agar (right). Bacterial stocks were streaked onto agar plates and incubated for 40 hours at 37°C.

2.3.2. Growth variations between subspecies of *S. equi* in different media

In order to compare growth of the industrial biomanufacturing strain (SE40327) against other *S. equi* strains a multi-well plate culture method was used. Growth of all strains in industrial fermentation media was compared to that of BHI, a nutrient rich media suitable for the growth of fastidious organisms. As can be seen in Figure 8, growth between the strains is not significantly different in fermentation media, although

SEE20561 reaches a lower maximum optical density. The error bars indicate that growth varies between biological replicates, an observation which has also been made at industrial fermentation culture volumes. Growth of all strains follows a similar trend until late log to stationary phase (8-10 hours), with SE40327 and SEZ20727 following the most similar growth curve.

In contrast, Figure 9 displays the growth profiles for the four strains in BHI. While variation between biological replicates appears to be greater in comparison to fermentation media, indicated by the larger error bars, the final OD reached for all four strains including SEE20561 are not significantly different. Stationary phase appears to be more stable in BHI media than in fermentation media, where small increases in OD are observed even up to 20 hours of growth. In addition, the time at which stationary phase is reached by each strain appears to vary more in BHI (9-13 hours). Surprisingly, SE40327 appears to be most similar in growth profile to SEZ20727 and SER17037, although overall the growth curves are similar for all strains.

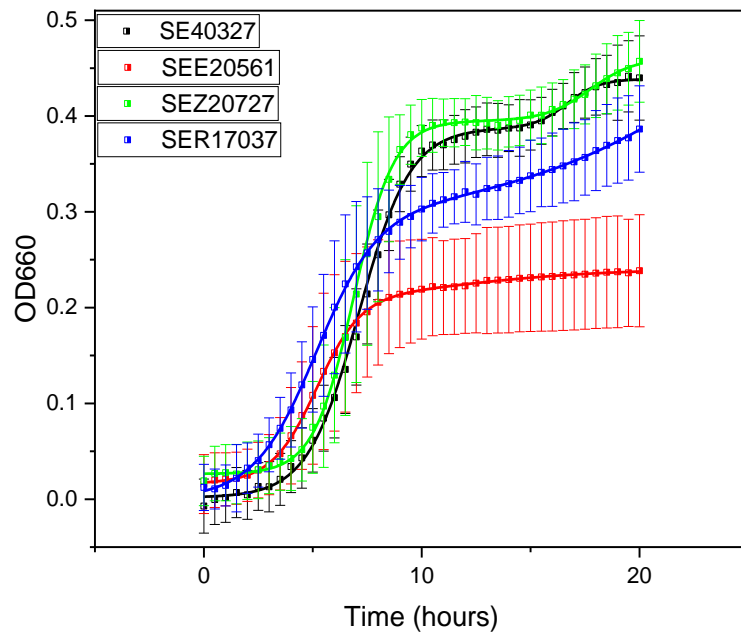


Figure 8 Comparison of growth over time of *S. equi* strains in fermentation medium. The strains represented are *S. equi* strain 40327 (SE40327), *S. equi* subsp *equi* DSMZ20561 (SEE20561), *S. equi* subsp *zooepidemicus* DSMZ20727 (SEZ20727) and *S. equi* subsp *ruminantium* DSMZ17037 (SER17037). The curves plotted are optical density at 660 nm over time. Points on the graph are representative of mean values over 3 biological replicates and error bars are standard deviation.

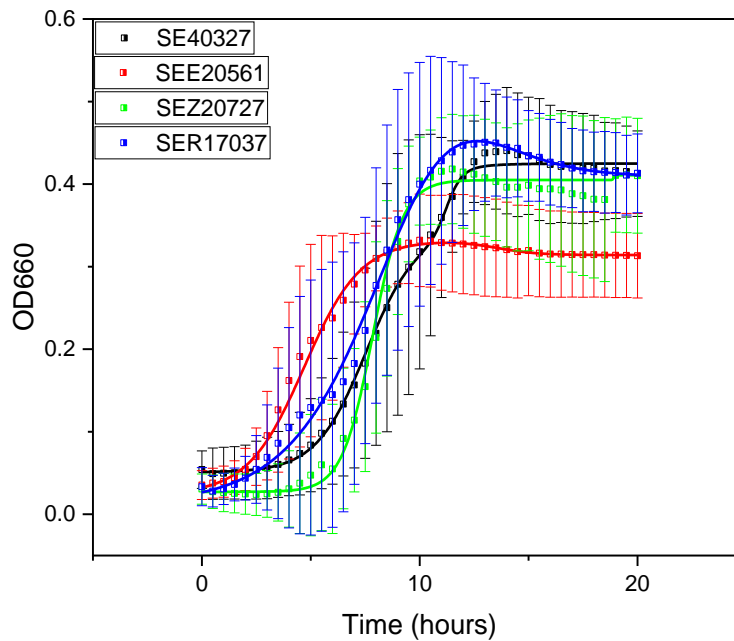


Figure 9 Comparison of growth over time in *S. equi* strains in BHI medium. The strains represented are *S. equi* strain 40327 (SE40327), *S. equi* subsp *equi* DSMZ20561 (SEE20561), *S. equi* subsp *zooepidemicus* DSMZ20727 (SEZ20727) and *S. equi* subsp *ruminatorum* DSMZ17037 (SER17037). The curves plotted are optical density at 660 nm over time. Points on the graph are representative of mean values over 3 biological replicates and error bars are standard deviation.

In order to observe variations in growth between media type growth curves were plotted of single strains in fermentation media and BHI media (Figures 10-13). It is apparent that no significant difference of the growth curves of SE40327 can be observed between the two media (Figure 10). This is in contrast to SEE20561 (Figure 11) in which a reduced lag time and increased maximum OD are visible, though the difference does not appear to be significant. Interestingly, SEE20561 reaches a lower average OD than SE40327 in both media types. Meanwhile the curves for SEZ20727 (Figure 12) appear to be very similar to SE40327 in that they do not differ significantly between media and similar lag time and stationary phase profile are observed. As in the growth curves above a notable small increase in OD is present during stationary phase for both SE40327 and SEZ20727 in fermentation media, whereas the curves in BHI level off in stationary phase in a more pronounced manner. For SER17037 (Figure 13), while no significant difference in final OD is observed between the two media the growth curves are different in shape. Lag time appears to be extended in BHI compared to fermentation media and larger error bars

indicate greater variation in growth between replicates. As with SE40327 and SEZ20727, stationary phase appears more distinct in BHI media with small increases in OD observable after initial stationary phase appears to have been reached. However, lag time for this strain appears to be significantly extended in BHI compared to fermentation media, which is not noted in the other strains. Overall, growth does not appear to be significantly different between fermentation media which is prepared specifically for industrial production of HA and BHI media which is recommended for the culture of fastidious pathogenic organisms.

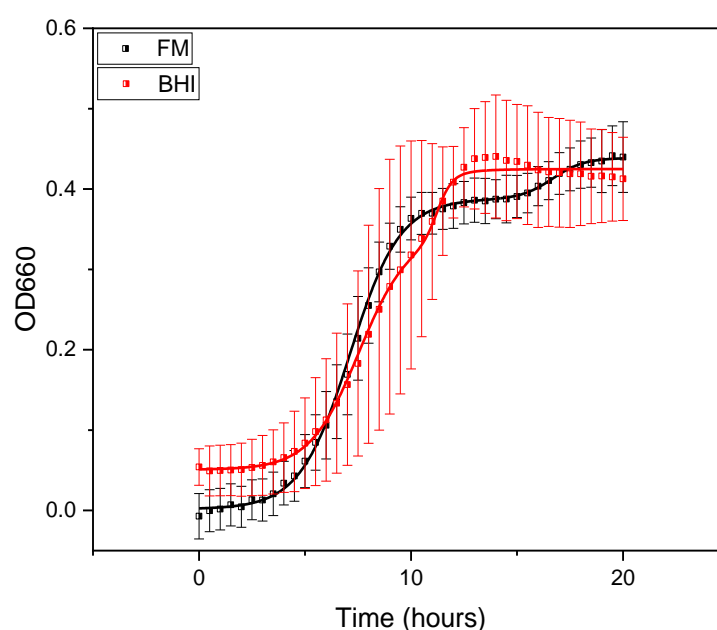


Figure 10 *S. equi* NCIMB 40327 growth in varying media. The curves plotted are optical density at 660 nm over time. Points on the graph are representative of mean values over 3 biological replicates and error bars are standard deviation. Data is taken from the SE40327 data in figures 8 and 9.

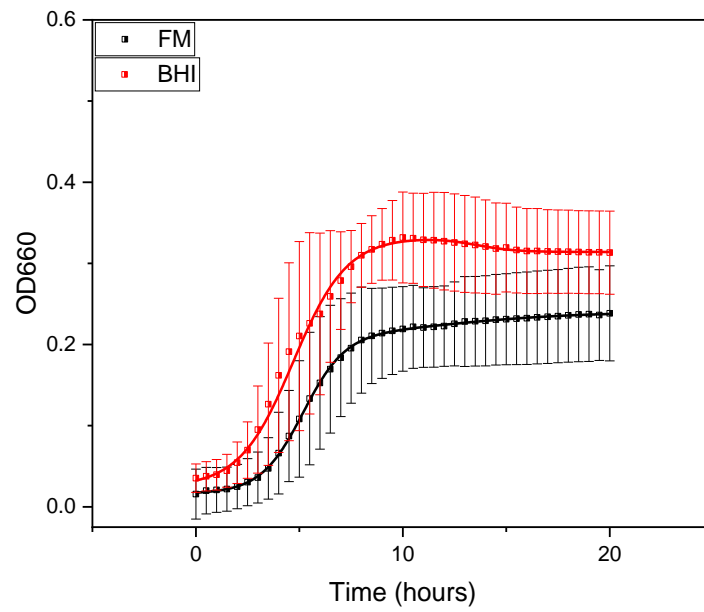


Figure 11 *S. equi subsp equi* (SEE20561) growth in varying media. The curves plotted are optical density at 660 nm over time. Points on the graph are representative of mean values over 3 biological replicates and error bars are standard deviation. Data is taken from the SEE20561 data in figures 8 and 9.

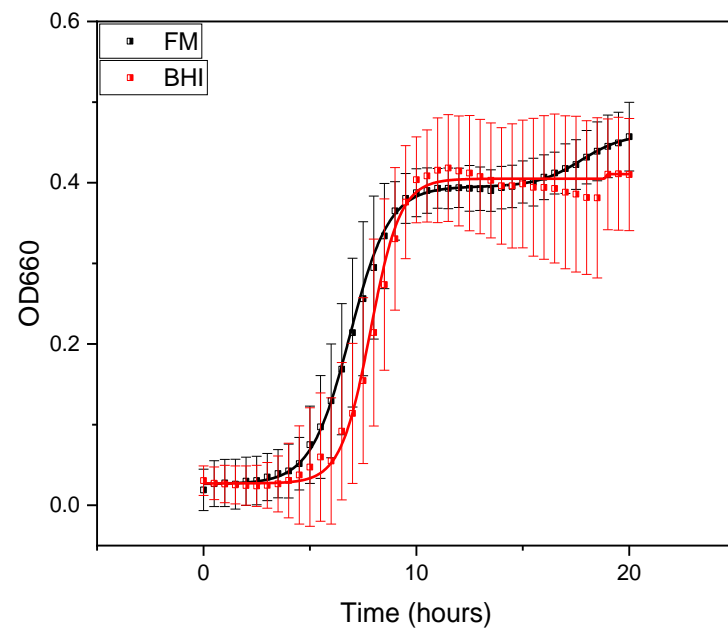


Figure 12 *S. equi subsp zooepidemicus* (SEZ20727) growth in varying media. The curves plotted are optical density at 660 nm over time. Points on the graph are representative of mean values over 3 biological replicates and error bars are standard deviation. Data is taken from the SEZ20727 data in figures 8 and 9.

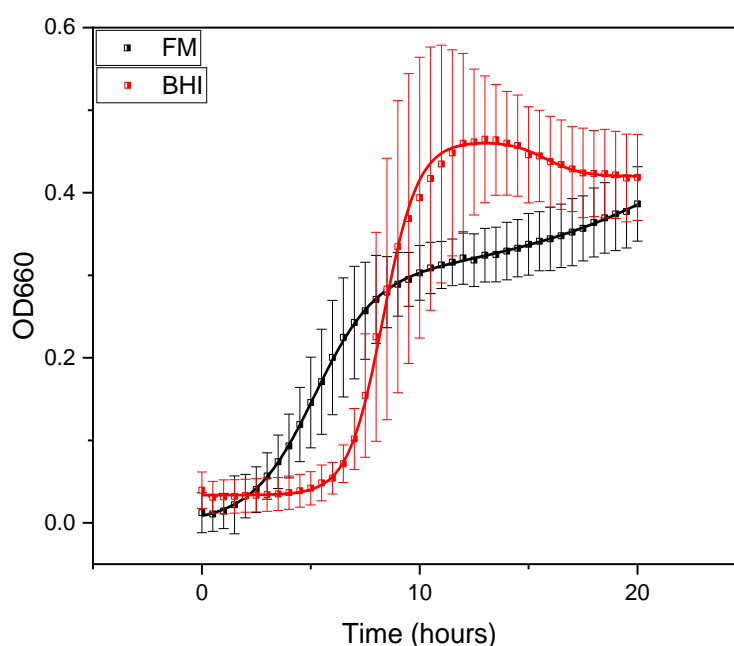


Figure 13 *S. equi subsp ruminatorum* (SER17037) growth in varying media. The curves plotted are optical density at 660 nm over time. Points on the graph are representative of mean values over 3 biological replicates and error bars are standard deviation. Data is taken from the SER17037 data in figures 8 and 9.

2.3.3. pH control and buffering

S. equi are lactic acid producing bacteria and industrial fermentation requires pH control to limit growth inhibition caused by a decline in pH. In order to assess the effects of acid production on the decline of pH of the culture medium, alteration in pH was assessed against growth over time, Figure 14 displays the results of these analyses. The starting pH of the fermentation media is 7, from the graph it is observed that following two hours of growth this drops to almost 6. Both growth and pH follow consistent trends over biological replicates. Overall an inverse relationship is evident between pH and growth, as was anticipated.

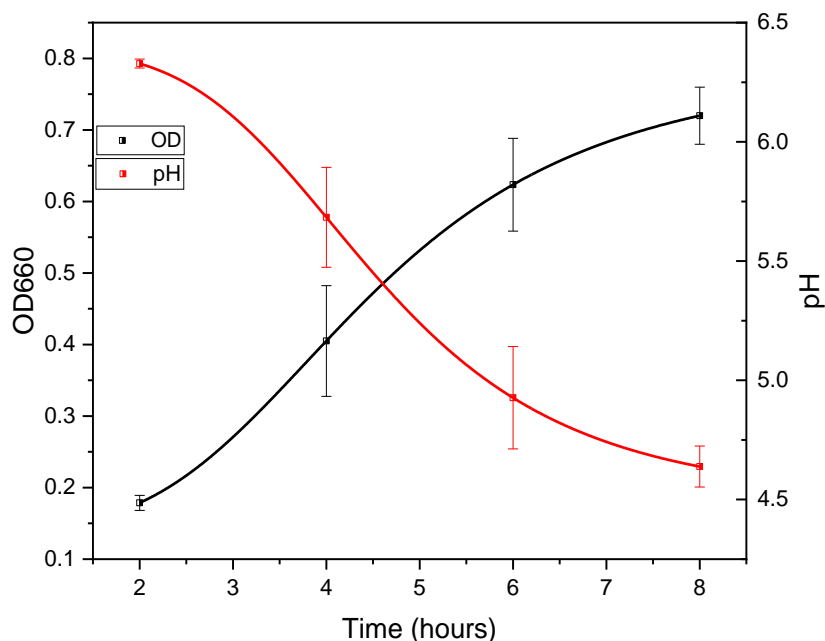


Figure 14 Growth vs pH over time of SE40327 culture in fermentation media. Optical density readings were performed at 660 nm. The points plotted represent mean over 3 biological replicates. Standard deviation was used for error bars.

The fermentation process which the small scale culture aims to mimic controls pH of the culture media by way of *in situ* pH monitoring and addition of 32% w/v sodium hydroxide solution (NaOH). In order to establish better pH control in small scale fermentation of HA, the effects of a series of widely-used biological buffers on growth of the bacteria was investigated. The buffers investigated were HEPES (effective pH range: 7.0-7.6, pH 7-7.2), Trizma base (effective pH range: 7-9, pH 10.5-12) and PIPES (effective pH range: 6.1- 7.5, pH 6).

Figure 15 displays the effect on growth of strain SE40327 of HEPES buffering of fermentation media at concentrations of 0, 10, 25 and 50 mM. While no significant difference on log phase growth can be observed between buffer concentrations, a significant difference in the final OD reached is seen between 0 and 50 mM HEPES concentrations. The increase is by approximately 0.1 OD, with the average OD at 20 hours of growth without buffering being 0.52 compared to 0.62 with 50 mM HEPES. With regards to the average maximum OD, buffering with 50 mM HEPES

increased this from 0.52 to 0.65. From the graph it is evident that 50 mM was the optimum HEPES concentration from those tested.

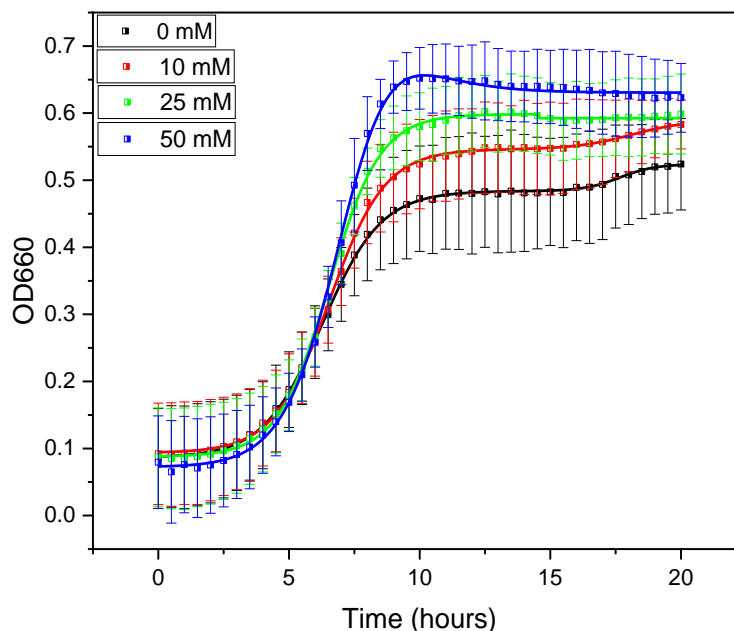


Figure 15 Effect of increasing HEPES buffer concentration on *S. equi* NCIMB 40327 growth. Growth was assessed in fermentation media (starting pH 7) with 0, 10, 25 or 50 mM HEPES buffer solution. The curves plotted are optical density at 660 nm over time. Points on the graph are representative of mean values over 3 biological replicates and error bars were calculated using standard deviation.

Following the results of the HEPES buffer concentration analysis, the growth of all four strains of interest was investigated in the maximum HEPES concentration tested (50 mM). From figure 16 an improvement in maximum OD is observed for all strains in comparison to culture in non-buffered fermentation media (Figure 8). It is also apparent that the strains display differences in both lag time and the time point at which stationary phase is reached. SEE20561 still does not reach as high a maximum OD as the other three strains and appears to have a slightly longer lag time than two of the strains. Whereas SE40327 appeared to display the longest lag time, yet reached a much higher final OD. SEZ20727 and SER17037 appear to have similar lag and log profiles, but SEZ20727 reaches a significantly higher OD upon reaching stationary phase. No significant difference between final OD is noted between SE40327 and SEZ20727

however. Whilst buffering with HEPES has increased the OD reached by all strains, it has emphasised differences in growth profiles between them which were not as evident in fermentation media without HEPES buffering or BHI media (Figures 8 and 9).

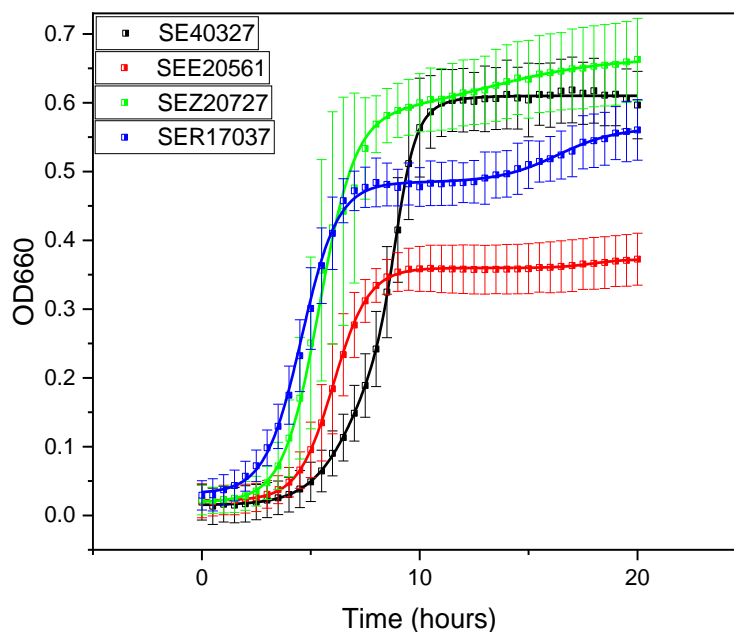


Figure 16 Comparison of strain growth in fermentation media supplemented with 50 mM HEPES buffer. Starting pH of the media was 7. The strains represented are *S. equi* strain 40327 (SE40327), *S. equi* subsp *equi* DSMZ20561 (SEE20561), *S. equi* subsp *zooepidemicus* DSMZ20727 (SEZ20727) and *S. equi* subsp *ruminantium* DSMZ17037 (SER17037). The curves plotted are optical density at 660 nm over time. Points on the graph are representative of mean values over 3 biological replicates and error bars were calculated using standard deviation.

The following four graphs (Figures 17-20) more clearly display the differences in growth profiles for each strain when cultured in buffered or non-buffered media. From all of the graphs it is clear that a significant increase in maximum OD occurs when the culture is buffered with HEPES. In the case of SEE20561, variation of growth between biological replicates appears reduced in buffered media, whereas for SEZ20727 the opposite is apparent. Only SEZ20727 and SER17037 experience a change in overall growth profile, with the lag time for both strains being slightly shorter in buffered media compared to non-buffered.

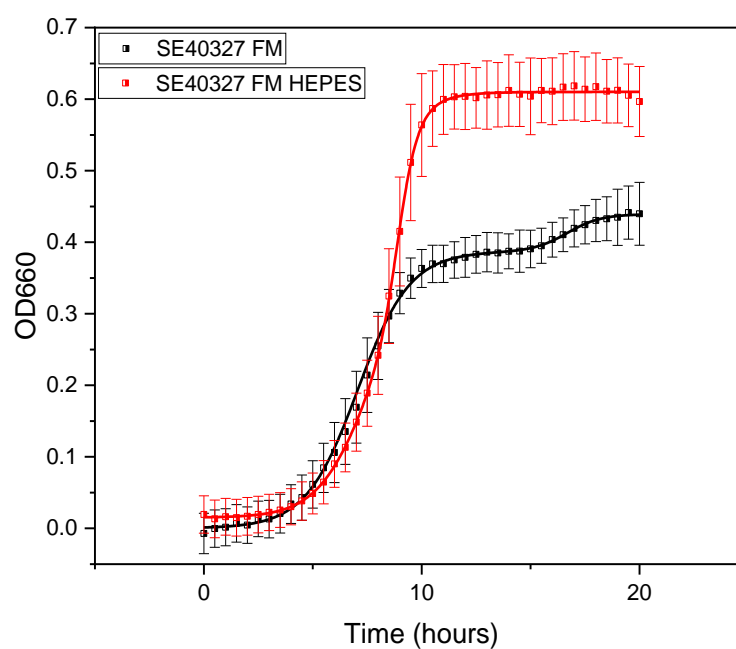


Figure 17 SE40327 growth in buffered and non-buffered fermentation. The graph displays growth of SE40327 presented as mean OD values versus time in hours at 37°C. Standard deviation was calculated for the error bars. Conditions presented are fermentation media only (SE40327 FM) and fermentation media supplemented with 50 mM HEPES solution (SE40327 FM). Data was taken from the SE40327 data in figures 8 and 16. The starting pH of both media was 7.

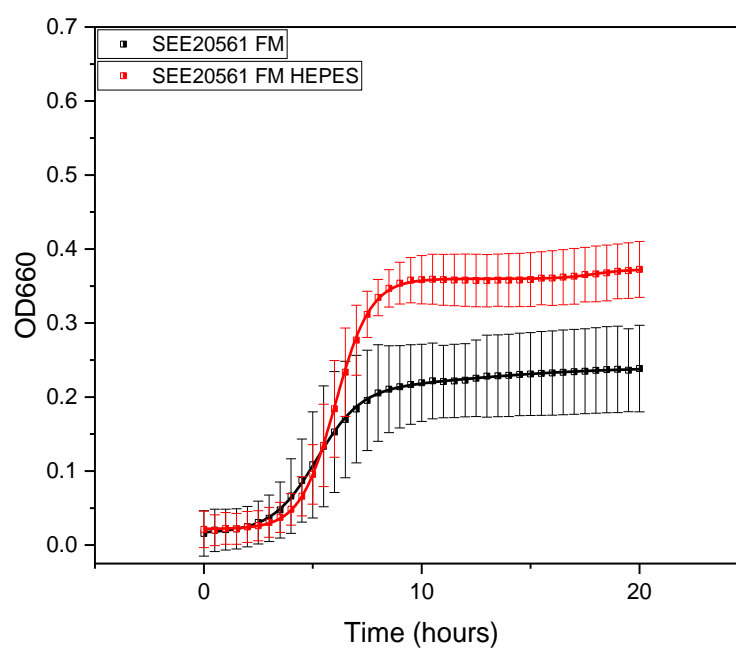


Figure 18 SEE20561 growth in buffered and non-buffered fermentation media. The graph displays growth of SEE20561 presented as mean OD values versus time in hours at 37°C. Standard

deviation was calculated for the error bars. Conditions presented are fermentation media only (SEE20561 FM) and fermentation media supplemented with 50 mM HEPES solution (SEE20561 FM HEPES). Data was taken from the SEE20561 data in figures 8 and 16. The starting pH of both media was 7.

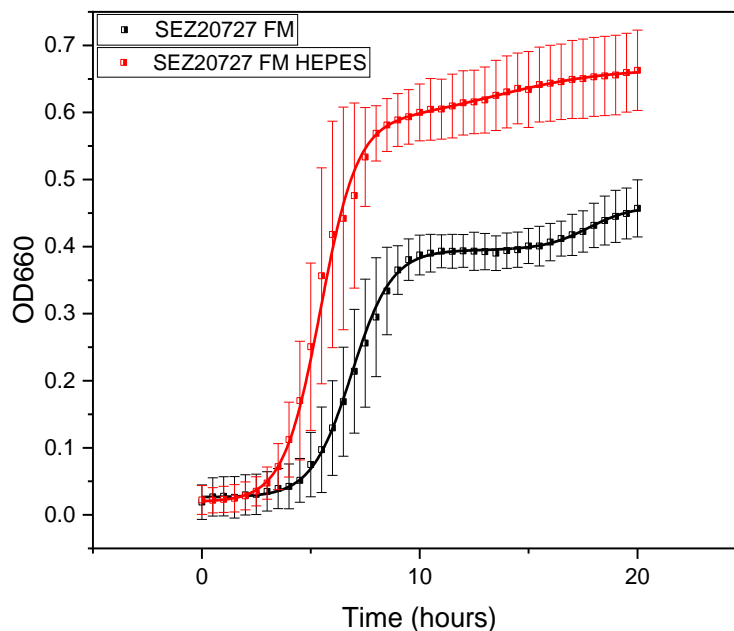


Figure 19 SEZ20727 growth in buffered and non-buffered fermentation media. The graph displays growth of SEZ20727 presented as mean OD values versus time in hours at 37°C. Standard deviation was calculated for the error bars. Conditions presented are fermentation media only (SEZ20727 FM) and fermentation media supplemented with 50 mM HEPES solution (SEZ20727 FM HEPES). Data was taken from the SEZ20727 data in figures 8 and 16. The starting pH of both media was 7.

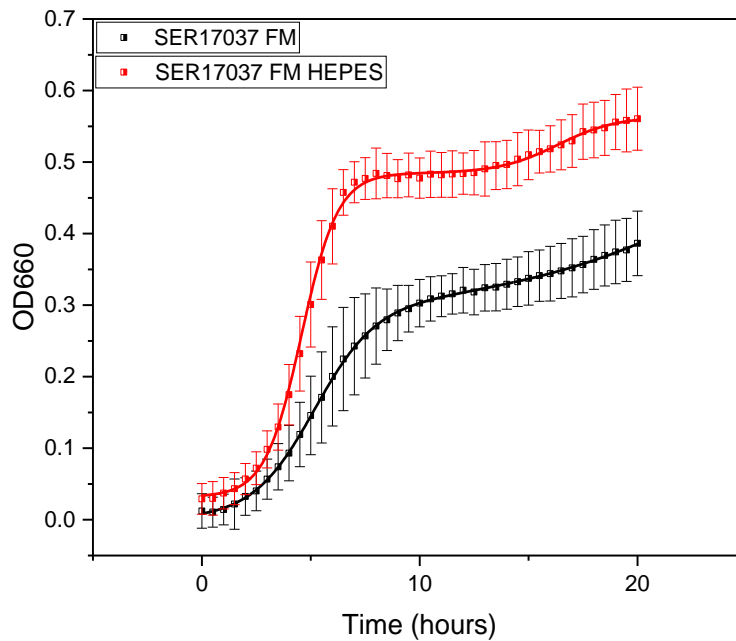


Figure 20 SER17037 growth in buffered and non-buffered fermentation media. The graph displays growth of SER17037 presented as mean OD values versus time in hours at 37°C. Standard deviation was calculated for the error bars. Conditions presented are fermentation media only (SER17037 FM) and fermentation media supplemented with 50 mM HEPES solution (SER17037 FM HEPES). Data was taken from the SER17037 data in figures 8 and 16. The starting pH of both media was 7.

Figure 21 displays the effect on growth of SE40327 with buffering at various concentrations of Tris buffer. Whilst an increase in final OD compared to non-buffered growth is observed at concentrations of 10 and 25 mM, the growth curves appear to become increasingly variable as Trizma concentration is increased. Furthermore, a concentration of 50 mM appears to have caused cessation of growth entirely.

The same concentrations of PIPES buffer were tested on growth of SE40327. No growth occurred in any concentration of PIPES, suggesting this buffer solution is inhibitory for this strain.

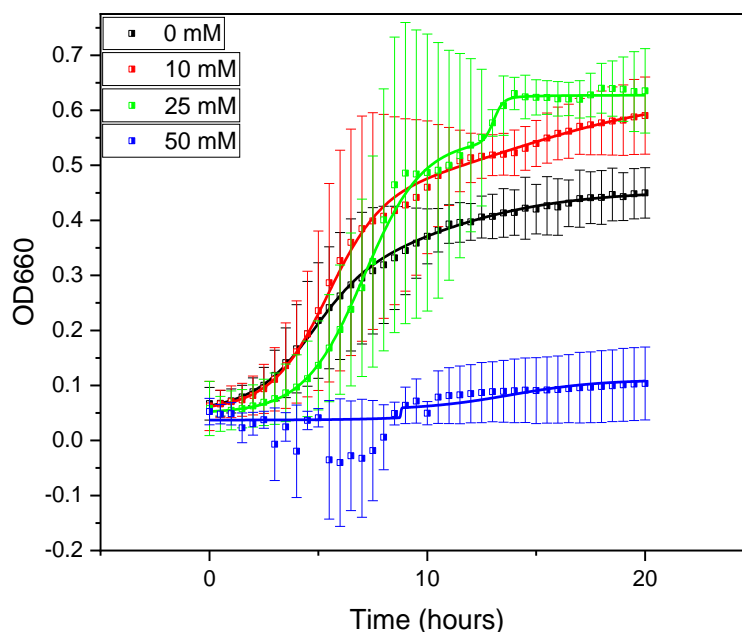


Figure 21 Effect of increasing Tris buffer concentration on *S. equi* NCIMB 40327 growth.

Growth was assessed in fermentation media with 0, 10, 25 or 50 mM Trizma base buffer solution.

The curves plotted are optical density at 660 nm over time. Points on the graph are representative of mean values over 3 biological replicates and error bars were calculated using standard deviation.

The starting pH of the media was 7.

2.3.4. Influence of media on capsule production of *S. equi* strains

The results of size exclusion HPLC analysis of filtered 8 hour culture samples are displayed in Tables 7 and 8. Only two of the strains were found to produce HA, with SEE20561 and SEZ20727 consistently failing to produce detectable concentrations. The concentrations from each biological replicate were corrected to the maximum OD achieved in order to normalise the data, as final OD and HA concentration varied substantially in some cases. The retention time consistently measured between 3.7 and 4 minutes which is within expected range for HA, whereas concentrations varied between approximately 60-200 ng/ μ L before correction for OD. For each strain the HA concentrations were analysed between the different media were analysed in Minitab 17 Statistical software. For SE40327, there was a significant difference between mean corrected HA concentrations for buffered and non-buffered fermentation media (One-Way ANOVA, $f = 7.03$, $df = 2, 6$, $p = 0.027$). However, HA

concentration in BHI media was not significantly different from either buffered or non-buffered fermentation media. In contrast for SER17037, HA concentrations were found to have no significant differences between all three media tested (One-Way ANOVA, $f = 3.58$, $df = 2; 6$, $p = 0.095$).

The variation in HA concentrations between strains was also analysed for each media type. No significant differences were found between strains in fermentation media (two-sample t-test, $t = -2.14$, $df = 2$, $p = 0.166$), BHI media (two-sample t-test, $t = -1.87$, $df = 4$, $p = 0.134$) or fermentation media buffered with HEPES (two-sample t-test, $t = -1.64$, $df = 2$, $p = 0.242$).

SE40327				
Media	Retention time (min)	HA concentration (ug/ml)	Final OD660	Corrected concentration to max OD (ug/ml)
FM	3.769	72.7	0.31	90.36
FM	3.892	89.46	0.35	100.2
FM	3.874	96.75	0.39	96.75
FM HEPES	3.809	60.15	0.19	209.22
FM HEPES	3.916	89.00	0.33	176.32
FM HEPES	3.9	125.14	0.66	125.14
BHI	3.893	97.80	0.36	107.16
BHI	3.913	90.06	0.29	122.95
BHI	3.961	92.73	0.39	92.73

Table 7. Hyaluronic acid concentration produced by SE40327. The table displays the results of HPLC analyses on cell stripped culture following 8 hours of growth in different media. Fermentation media (FM), fermentation media buffered with 50 mM HEPES buffer (FM HEPES) and Brain Heart Infusion media (BHI). Final optical density values are displayed along with the measured hyaluronic acid concentration and retention time for each analysed sample. Concentrations were corrected to the maximum OD reached for each media. Each entry represents a single biological run of 15 pooled technical replicates (wells).

SER17037				
Media	Retention time (min)	HA concentration (ug/ml)	Final OD660	Corrected concentration to max OD (ug/ml)
FM	3.797	65.7	0.11	200.23
FM	3.867	96.60	0.24	133.26
FM	3.876	117.53	0.32	117.53
FM HEPES	3.783	58.96	0.073	387.26
FM HEPES	3.94	110.37	0.24	224.12
FM HEPES	3.862	207.12	0.48	207.12
BHI	3.932	131.15	0.36	141.52
BHI	3.92	76.00	0.16	183.3
BHI	3.947	103.14	0.39	103.14

Table 8. Hyaluronic acid concentration produced by SER17037. The table displays the results of HPLC analyses on cell stripped culture following 8 hours of growth in different media. Fermentation media (FM), fermentation media buffered with 50 mM HEPES buffer (FM HEPES) and Brain Heart Infusion media (BHI). Final optical density values are displayed along with the measured hyaluronic acid concentration and retention time for each analysed sample.

Concentrations were corrected to the maximum OD reached for each media. Each entry represents a single biological run of 15 pooled technical replicates (wells).

2.4. Discussion and Conclusions

2.4.1. Sulphur limited media and its effect on HA production

From the original patent submission (WO1992008799A1), the use of media containing a limited sulphur supply was described as an attempt to limit the growth rate of these bacteria without inhibiting HA production. In that patent, it is also stated that limiting other crucial components such as carbon, nitrogen, sodium, phosphorus, potassium, magnesium, iron, manganese and zinc as an alternative to sulphur was acceptable. Since no explanation or reason was given for requiring this limitation of growth during colony screening or inoculum preparation, this might be considered an attempt to divert the metabolism of glucose away from biomass and towards HA production.

For the experimentation described, the sulphur limited agar plates were acquired from the same source as used by the company and the same sulphur limited broth recipe was used in order to replicate the manufacturing process as closely as possible. Whether limiting sulphur in the growth media is an effective method of promoting HA production is questionable from the colony morphology of the strains when cultured on sulphur limited agar in comparison to BHI agar plates, with colonies appearing exceptionally more mucoid on BHI. This suggests that the use of sulphur limited agar/broth could have an opposing effect to the one intended. In addition to this, the results of the HA concentration analyses presented in Tables 7 and 8 shows that HA production does not differ significantly between fermentation medium and BHI medium. The most obvious distinct feature of the fermentation medium is the high glucose concentration of 74 g/L (411 mM). Previously it has been suggested that glucose is one of the limiting nutrients in HA fermentation, making the lack of significant difference in HA yield between such a high concentration in fermentation media and the 2 g/L concentration of BHI an unexpected result. Of course neither media are chemically defined, thus the role of other unknown factor(s) being at play is possible as the vitamin,

nitrogen, amino acid, and carbon concentrations within yeast extract, calves brain and beef heart preparations are variable. As such the isolated effect of any one component of these extracts cannot be observed. As previously mentioned a number of studies have been conducted to investigate the effects of variation to culture parameters on HA fermentation in *Streptococcus*. There have also been attempts to identify vital media components for growth and HA production. The development of a chemically defined media (CDM) enabled the identification of 11 amino acids which were essential for growth of *Streptococcus equi* subsp. *zooepidemicus* ATCC 35246 (Armstrong, Cooney and Johns, 1997). The CDM developed did not significantly increase HA production compared to a complex media containing yeast extract, however it did result in a significant increase in HA molecular weight. Another study investigated the impact of variation of mineral ion concentrations in media on growth and HA production in *Streptococcus equi* subsp. *zooepidemicus* ATCC 39920 (Pires, Eguchi and Santana, 2010). The results showed that absence of sodium ions (Na^+) alone induced an increased HA production of 22% compared to controls. However the addition of Na^+ to the media increased the molecular weight of the HA produced by over 20%, suggesting that production and molecular weight of HA are enhanced by opposing conditions.

However, glucose is still a media component to consider when developing an optimised system for HA production, with the reported concentrations used being 10-60 g/L (Chong *et al.*, 2005). An investigation into the effects of initial glucose concentration (IGC) on HA production revealed that whilst bacterial growth was unaffected by its presence or absence, the highest molecular weight and lowest concentration of HA was achieved in complete absence of glucose. In these conditions metabolism switched to mixed acid and lactate production was reduced (Pires and Santana, 2010). HA molecular weight was lowest with the introduction of glucose but gradually increased between concentrations of 5-45 g/L, with 25 g/L being optimum for HA synthesis over cell growth from the glucose supply. HA concentration (g/L) remained consistent with increasing glucose concentrations 5-45 g/L. In another investigation the IGC at which both

HA concentration and molecular weight were highest was found to be 40 g/L with a further increase of 50-60 g/L being inhibitory for production (Don and Shoparwe, 2010). This was speculated to be due to cells reacting to increased osmotic pressure at 40 g/L compared to 10-30 g/L and the occurrence of interference from competing pathways such as lactic acid and peptidoglycan production at glucose concentrations above 50 g/L. In addition to concentration, the source of glucose is a consideration to be made in the preparation of HA fermentation media. During the development of a serum free medium for growth and HA production by *S. zooepidemicus* NJUST01 the use of starch as the carbon source in place of glucose was found to be optimum for HA production. This result was associated with the observation that a slower and less severe decline in pH occurred when starch was used as the glucose source, which was speculated to be due to a reduction in lactic acid production (Zhang *et al.*, 2006). A re-examination of the current fermentation media components and the influence of each on HA production could provide beneficial information about the effectiveness of the recipe. Whilst the development of a CDM may enhance HA production and molecular weight in this strain, as previously mentioned it is considered not cost effective for industrial scale production of HA.

2.4.2. Selecting an appropriate buffer for pH control in small scale fermentation

Under manufacturing conditions HA fermentation is performed in volumes of 80-130 L. For research purposes such volumes are unnecessary and wasteful of resources, thus a protocol for 'small scale fermentation' was developed. The use of 96 well plates was preferred to shake flask, mainly due to the lower volume requirements, increased ease of optical density monitoring in a plate reader and the ability to analyse more than one media/strain/antibiotic concentration at a time.

Streptococci are lactic acid producing bacteria, a characteristic which is employed as an offensive strategy against competing bacteria in the environment. In *in vitro* culture conditions however, lactic acid production may rapidly decrease the pH of the culture media and autoacidification rather than nutrient depletion is a common cause of premature cessation of

growth of lactic acid producing bacteria (Papadimitriou *et al.*, 2016). The small increases in OD observed in all strains during stationary phase when cultured in fermentation media could possibly indicate that attenuation of growth is not a result of exhaustion of nutrients but of autoacidification, highlighting the importance of pH control. When performing fermentation at large volumes in chemostats it is possible to continuously monitor and maintain the culture pH. In the case of the manufacturing process in question this is done by controlled addition of 16% sodium hydroxide solution (NaOH). Unfortunately this method of pH control is problematic outwith chemostat operation and so for multi-well plate culture it was more suitable to identify an appropriate buffering agent to reduce the effects of autoacidification.

The starting pH of the fermentation media was pH 7. Therefore buffers for which included this pH in the effective range were selected for trial purposes. The buffers were HEPES, PIPES and Tris which are widely used with biological systems. HEPES is a widely used buffering agent for the culture of a variety of organisms; PIPES buffer is an organic buffer with a useful pH range of 6.1-7.5 and Trizma base buffer is commonly used in cell culture media and has an effective pH range of 7-9. Therefore all were considered as suitable candidates for pH control of the cultures in 96 well plates.

HEPES has a stated effective range of pH 6.8 – 8.2 with the fermentation media at pH 7 falling at the lower end of this. The recommended concentration range for growth media is 10 – 25 mM, although concentrations as high as 100 mM have been used for other *Streptococcus* species (Tremblay *et al.*, 2009) so growth at 50 mM was also investigated. Rather than negatively affecting growth, 50 mM HEPES concentration provided the greatest benefit of the three concentrations tested (10, 25 and 50 mM), suggesting it is not detrimental to growth of this strain. Although a greater maximum OD was reached when media was buffered with HEPES compared to non-buffered media, log phase did not appear to be significantly extended, suggesting that either autoacidification still occurs although at higher cell density and/or that nutrient deprivation occurs at the same time point regardless of the rate of pH decline. PIPES buffer, as

mentioned, is an organic buffer with a useful pH range of 6.1-7.5. The pH of the fermentation media falls more centrally within this range compared to the range of HEPES and so PIPES was considered a potentially more suitable candidate for pH control of the cultures. Having previously been investigated in *S. thermophilus* (Somkuti and Gilbreth, 2007) it was expected that this buffer would also be suitable for growth of the strain of interest. Unfortunately this was not the case and growth of *S. equi* was completely inhibited under conditions tested (data not shown). This was an unexpected result considering that previously PIPES has been used at concentrations as high as 100 mM for culture of other *Streptococcus* species (Dierksen, Ragland and Tagg, 2000), but unfortunately it does not appear to be suitable for SE40327 (and presumably other *S. equi*) at the ranges tested in a 96 well plate format. Tris buffer is commonly used in cell culture media and was readily available to be tested in the strain of interest, the useful pH range of Trizma base is 7-9. Whilst an increase in final OD was observed for concentrations of 10 and 25 mM Trizma base compared to non-buffered media the effect was not significantly different to that seen at the same concentrations of HEPES buffer. In addition Trizma appeared to have a detrimental effect on growth of *S. equi* at a concentration of 50 mM. This contrasts to 50 mM HEPES which produced the greatest increase in final OD in this strain. Due to the fact the useful range of HEPES was more suitable to the fermentation media at pH 7, it was chosen over Trizma base as the more suitable candidate for pH control in this study.

Currently pH control is carried out at production scale by the addition of NaOH. Whilst NaOH addition alone is likely more cost effective (approximately £83 for 5L of 32% NaOH solution versus approximately £600 for 4L HEPES) than HEPES buffering, the use of NaOH may contribute to environmental pollution when it is discarded within waste fermentation media and demonstrates a significant hazard should large quantities be introduced into the environment (EnvironmentAgency, 2018). The addition of HEPES or another suitable buffering agent to the fermentation media may allow a reduction in the quantity of NaOH used to maintain steady pH during a fermentation cycle. Even a change of 0.5 of

pH level was found to decrease the molecular weight and concentration of HA produced during fermentation of *S. equi* ATCC 6580 (Kim *et al.*, 1996) so even fluctuations in pH may have influence over product properties.

2.4.3. Hyaluronic acid production in *S. equi* strains

As SE40327 is a strain of *S. equi* subsp. *equi* it was expected that of the three *S. equi* strains selected for comparison it would most closely resemble *S. equi* subsp. *equi* 20561 (SEE20561) in growth profile and colony morphology. This was not revealed to be the case, with SEZ20727 being the closest in growth profiles in both fermentation and BHI media. All strains however were very similar, with only SEE20561 deviating significantly from the rest in growth profile. Previously, both SEE20561 and SEZ20727 were described as encapsulated (Speck *et al.*, 2008), however HPLC analysis for HA did not reveal detectable concentrations in either of these two strains under the growth conditions investigated, nor did their colony morphology appear mucoid on sulphur limited or horse blood agar in comparison to SE40327 and SER17037. SER17037 was also previously described as mucoid when it was first isolated in 2004 (Fernandez *et al.*, 2004), however its capability as a HA producer has never before been investigated. From the results of the HPLC analysis it is clear that SER17037 is just as, if not more, capable of producing appropriate concentrations of HA than the currently favoured strain. Whilst the *p*-values from the statistical analyses of the results indicated that no significant difference exists between HA concentrations produced by each strain in the same media, it is clear from the data that SER17037 is capable of producing higher concentrations of HA in comparison to SE40327 under the conditions tested, typically in excess of 100 ng/μL. In addition, the increase in average concentration/yield of HA observed in buffered fermentation medium compared to non-buffered is likely a result of increased cell numbers prior to the beginning of stationary phase as indicated by the greater maximum OD.

From the data presented and the similarity of HA concentration between fermentation media and BHI media it is proposed that the glucose concentration in the current fermentation media recipe could be reduced without detriment to HA yield, as average concentration is not significantly

increased compared to BHI media which contains a much lower concentration of glucose. In addition it may be beneficial to investigate the batch to batch variation in the media as components like yeast extract are not chemically defined and variations in trace elements, amino acids and vitamins will occur. This knowledge would provide the first step towards developing a more defined media for HA production with this strain. The caveat to undertaking this work is that development of CDM is time consuming and as previously mentioned not cost effective for large scale production. However a greater knowledge of the components of the media may be useful for future. With regards to pH control, whilst the use of buffers is a suitable alternative to NaOH addition for a small scale model of fermentation, whether its use in manufacturing scale culture would be cost effective and necessary is uncertain. *S. equi* subsp. *ruminantium* has been previously characterized in terms of biochemical and sequence analyses (Fernandez *et al.*, 2004; Speck *et al.*, 2008). However its capability as a HA producer has not yet been investigated despite its close relation to other *S. equi* subspecies which are commonly used for HA fermentation (Boeriu *et al.*, 2013). From the data presented it can be concluded that this strain is capable of producing similar concentrations of HA than a currently used industrial strain and may be a candidate strain for HA production in the future.

3. Comparative Genomics of *S. equi* subspecies

3.1. Introduction

The *Streptococcus equi* strain SE40327 has been in use for the industrial manufacture of HA containing medical devices for the last 30 years. In that time genome sequencing has become a useful resource in the analysis of bacterial processes. To this point, the genome of this strain had not been sequenced, nor had that of the reference strains selected for this study *S. equi* subsp. *equi* 20561 (SEE20561) and *S. equi* subsp. *zooepidemicus* 20727 (SEZ20727). Thus the project aimed to extract the DNA of these strains for genome sequencing and analysis. The genome of *S. equi* subsp. *ruminatorum* 17037 (SER17037) has previously been sequenced and is available in NCBI under the strain number CECT 5772 (Genbank accession GCA_000706805.1). In addition, mucoid and non-mucoid phenotypes were identified for SE40327 which did not revert following repeated passages on agar (see Chapter 2) both variants were to be submitted for sequencing in an attempt to identify any mutations which have occurred which may have an effect on HA production.

As noted previously, HA biosynthesis is conferred by multiple genes encoded in one or two operons in *S. equi* subsp. *zooepidemicus* and *S. equi* subsp. *equi* respectively. In addition, growth phase-dependent production of hyaluronidase can play a role in HA production. HA biosynthetic genes and hyaluronidase are the main focus of this investigation although selected other notable genomic differences are considered.

It has been shown that the *has* operons of *S. equi* subsp. *equi* and *S. equi* subsp. *zooepidemicus* species were different in that all five genes required for HA production, hyaluronan synthase (*hasA*), UDP-glucose dehydrogenase (*hasB*), UDP-glucose pyrophosphorylase (*hasC/galU*), (*hasD/glmU*) and phosphoglucoisomerase (*hasE/pgi*) were contained within the *zooepidemicus* operon, whereas in *equi* a separate operon contains *hasD* and *hasE* (Blank, Hugenholtz and Nielsen, 2008). The structure of this operon in SER17037 has not previously been investigated and is unknown for SE40327. Therefore confirmation of the operon structure in these two strains will aid in the investigation into capsule production within these strains. In addition to this

any differences observed between the sequences of the *has* operon genes between strains may provide avenues to follow.

It has previously been suggested by Speck et al. (2008) that SER17037 is hyaluronidase negative whereas SEE20561 and SEZ20727 are hyaluronidase positive. However the presence or absence of hyaluronidase genes within SE40327 is unknown. Sequencing of the genomes of this strain and the reference strains will aid in the elucidation of hyaluronidase genes, whether functional or not.

Finally, sequencing of the genome of the strains will facilitate the design of molecular experiments which will be discussed in chapter 4 and will provide a more accurate database for proteomic investigations discussed in chapter 5. Seven strains were submitted for sequencing, these were a working stock and master stock of SE40327 (40327W and 40327M), the original strain submitted to NCIMB at time of patent (30100), SEE20561, SEZ20727 and the mucoid and non-mucoid phenotypes isolated in one of the stocks of 40327 (40327MV and 40327NMV). The genome of SER17037 was already available.

3.2. Materials and Methods

3.2.1. DNA extraction and sequencing

DNA extraction of samples was carried out using the MasterPure™ DNA Purification Kit (Epicentre). Bacteria were grown on sulphur limited agar plates at 37°C for 48 hours. Single colonies were inoculated into 50 mL of 37°C BHI media and incubated over night with 150 RPM shaking. Mucoid strains were further incubated for 30 minutes following addition of 1x volume of 0.2 mg/mL hyaluronidase (Sigma-aldrich) in BHI to degrade the capsule. Bacteria were pelleted at 4,969 x g for 10 minutes and the pellets were washed 3 times with 500 µL PBS. All liquid was removed from the pellets which were then resuspended in 300 µL of Tissue and Cell lysis solution containing 1 µL of Proteinase K (20mg/mL) (Qiagen). The samples were incubated at 65°C for 15 minutes and vortexed briefly 3 times during this time. Samples were left to cool on the bench for 15 minutes before 1 µL of RNase (Merck) was added and samples incubated at 37°C for 30 minutes. Protein was removed by adding 175 µL of MPC Protein Precipitation Reagent, vortexing for approximately 10 seconds and

centrifuging at 4°C for 10 minutes at 10,000 x g. The supernatant was transferred to a clean Eppendorf tube and the pellets were discarded. To the new samples, 500 µL of isopropanol was added and the tubes were inverted several times to mix thoroughly. The samples were centrifuged at 4°C for 10 minutes at 15,000 x g and the isopropanol carefully poured away. The pellets were washed twice with ice cold 70% ethanol with brief centrifugation to prevent the pellet from dislodging. The ethanol was removed and the tubes left open and horizontal for 5 minutes to dry the pellets. The DNA was then resuspended in 60 µL of HyClone™ Molecular Grade Water (Fisher Scientific). DNA concentration and purity was measured on NanoDrop 2000. A minimum 260/280 ratio of 1.8 was considered acceptable. Samples were stored at -80 °C until submitted to Edinburgh Genomics for sequencing via Illumina paired-end sequencing.

3.2.2. Genome assembly, annotation, species comparison and MLST typing

Sequence data was assembled into contigs using CLC Genomics Workbench 9.0 (<https://www.qiagenbioinformatics.com/>). The sequences were then uploaded to the RAST annotation server (<http://rast.nmpdr.org/>) for gene annotation. The RAST SEED Viewer was used to compare genomes at sequence level.

3.2.3. Analysis of hyaluronidase genes of each *S. equi* species

The sequences of genes annotated by RAST as “Hyaluronate lyase (phage associated)”, “Hyaluronoglucosaminidase”, “Hyaluronate lyase precursor” and “Phage hyaluronidase” were blasted using NCBI BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>), InterPro (<https://www.ebi.ac.uk/interpro/>) and Pfam (<https://pfam.xfam.org/>) in order to confirm the annotation. Only RAST annotated genes which also matched to Hyaluronidase protein families in NCBI, InterPro and Pfam were considered true hyaluronidase enzymes. In addition, genes which matched to beta-N-acetylglucosaminidase in the above databases were separately considered for further analyses. Gene protein and nucleotide sequences were compared using Multalin (<http://multalin.toulouse.inra.fr/multalin/>). Phylogenetic trees were constructed using MAFFT (<https://mafft.cbrc.jp/alignment/server/>).

3.2.4. Identification of hyaluronic acid associated genes in strains of interest

The genome of reference strain *S. equi* subsp. *equi* 4047 has been completely sequenced and is available in NCBI (accession number FM204883.1), this was uploaded to RAST in order to obtain a RAST annotated spreadsheet of this genome. The sequences pertaining to the *has* operon were identified within the RAST annotation by searching for the amino acid sequence of SEQ_0269 (hyaluronan synthase) identified in NCBI (Steward, Robinson and Waller, 2016). From this position the other capsule production associated genes downstream were identified in the RAST annotation. The amino acid sequences of the identified genes in *S. equi* subsp. *equi* 4047 were then used to search for and isolate the corresponding gene in the sequenced strains of interest. Promoter systems were identified using BPROM (<http://www.softberry.com/berry.phtml?topic=bprom&group=programs&subgroup=gfindb>).

3.2.5. Comparison of genes between strain phenotypes

The RAST annotations of the genomes of mucoid and non-mucoid variants of NCIMB 40327 were compared using the sequence based comparison tool in the RAST SEED viewer. The table was exported to Microsoft excel and genes sorted by their percentage identity between sequences. Genes absent in one or other variants genome sequence were identified. Genes known to be involved in HA synthesis were examined for any differences in sequence identity between mucoid and non-mucoid phenotype. Sequences were compared using Multalin (<http://multalin.toulouse.inra.fr/>).

3.2.6. Comparison of *hasC* paralogs

The paralogs of *hasC* were identified in the RAST annotated genomes of interest using the sequences from the type strain *S. equi* subsp. *equi* 4047 where possible. The protein and nucleotide sequences were compared using Multalin.

3.3. Results

Genome information of the sequenced strains is presented in table 9. Additional strains obtained from NCBI are included for comparison.

Strain	genome size (bp)	no. contigs	GC content (%)
30100	2,128,502	183	41
40327M	2,146,851	240	41
40327W	2,131,047	213	41
SEE20561	2,621,643	1171	44.8
SEZ20727	2,078,314	194	41.5
40327MV	2,156,647	259	40.7
40327NMV	2,210,074	380	41.8
SER17037	2,138,288	117	41.4
SE4047	2,253,793	1	41.3
SEZ35246	2,167,264	1	41.6

Table 9 Genome information of sequenced strains. The information presented within the table includes the genome size in number of base pairs (bp), the number of contigs assembled for each genome and the guanine and cytosine content of the genome (GC%). Shaded boxes indicate strains which were not sequenced in this project for comparison. These genomes are available on NCBI (<https://www.ncbi.nlm.nih.gov/>).

3.3.1. Hyaluronidase gene distribution across pathogenic Group C Streptococci species

For SE40327 four genes were annotated by RAST as Hyaluronate lyase (phage associated), one was identified as Hyaluronoglucosaminidase, two as Hyaluronate lyase precursors and four as Phage Hyaluronidases. Two of the “hyaluronate lyase (phage associated)” genes matched to bacterial hyaluronidase (IPR009860) in Interpro and hyaluronidase 1 (HylP) in Pfam (PF07212). They also matched in NCBI Blast to *S. equi* hyaluronoglucosaminidase and so for the purpose of this study were designated SE40327Hyal1 and SE40327Hyal2. The RAST annotated hyaluronoglucosaminidase gene matched in both Interpro and Pfam to beta-N-acetylglucosaminidase and also matched to hyaluronidase in NCBI Blast, this was designated SE40327bNAG. The other genes did not match to hyaluronidases or beta-N-acetylglucosaminidases in the databases mentioned. A comparison of the protein sequences of the three selected genes is shown in figure 22.

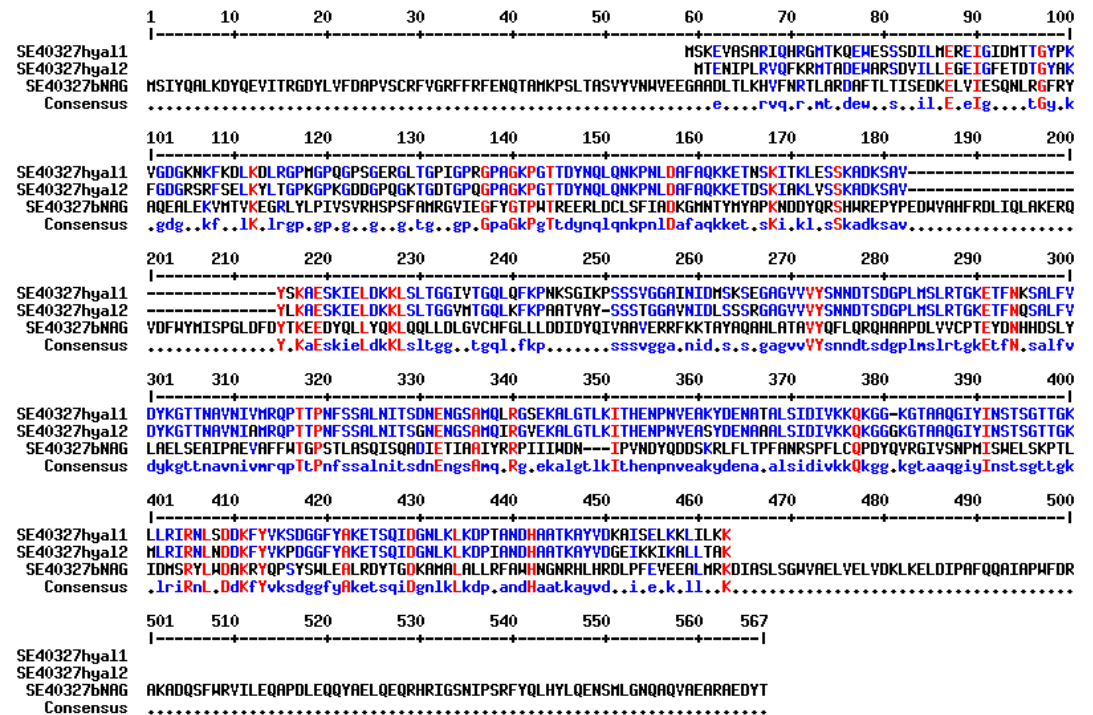


Figure 22 Comparison of the protein sequences of hyaluronidase and beta-N-Acetylglucosaminidase genes of SE40327. Sequences were identified and confirmed as described using the RAST annotated genome. Comparison was carried out using Multalin software (<http://multalin.toulouse.inra.fr/multalin/>). Sequence variation is observed between all genes and SE40327bNAG appears to be longer in sequence than the other genes.

S. equi subsp *equi* 4047 is a strain which was identified as the cause of strangles in a horse in England in 1990 (Kelly *et al.*, 2006). The whole genome of this strain is available from NCBI (accession FM204883). This was submitted to RAST for annotation in order to compare sequences to SE40327. As with SE40327, four genes were annotated as Hyaluronate lyase (phages associated), one as Hyaluronoglucosaminidase, two as Hyaluronate lyase precursors and four as Phage Hyaluronidases. One of the Hyaluronate lyase (phage associated) genes matched to hyaluronidase and the Hyaluronoglucosaminidase gene matched to beta-N-acetylglucosaminidase in both Interpro and Pfam. Both genes also matched identity in NCBI Blast to hyaluronidase and hyaluronoglucosaminidase. They were designated SE4047Hyal1 and SE4047bNAG and were compared to the three genes identified from SE40327. SE40327Hyal1 was identical at both protein and nucleotide level to the SE4047Hyal1. However SE40327Hyal2 displayed protein

sequence variation from both SE40327Hyal1 and SE4047Hyal1 (fig 23). The most notable feature between the sequences is a series of sequence variation between amino acid positions 1-80. In NCBI the section of protein sequence annotated as hyaluronidase of the genes is amino acids 94-372. Within this region sequence similarity is greater, with two regions of divergence at positions 154-160 and 361-373.

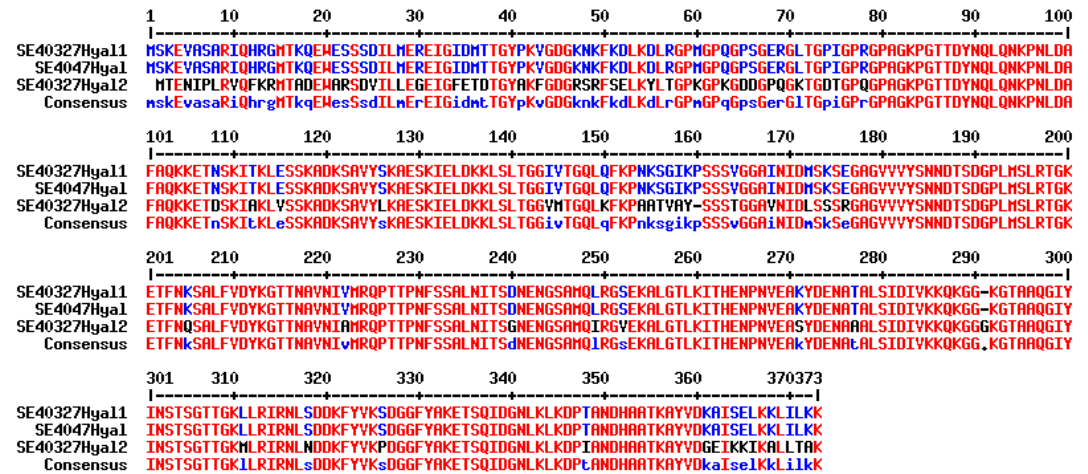


Figure 23 Hyaluronidase protein sequences of SE40327 compared to the sequence of SE4047.

Two hyaluronidase sequences obtained from the RAST annotated genome sequence of SE40327 (SE40327Hyal1 and SE40327Hyal2) were compared to the sequence from SE4047 (SE4047Hyal1). Sequences were compared using Multalin software (<http://multalin.toulouse.inra.fr/multalin/>). Variations in sequence are observed throughout between Hyal1 and Hyal2 genes, with larger stretches of variation observed at both protein terminals.

The protein sequence of SE40327bNAG and SE4047bNAG were found to be 100% identical, as were the nucleotide sequences (data not shown). Upon comparing to two strains of *S. equi* subsp. *zooepidemicus*, SEZ20727 and ATCC 35246 a series of single amino acid substitutions throughout the sequence were observed. Some of these single amino acid differences were observed between SE40327 and one of the SEZ strains but not the other (fig 24).

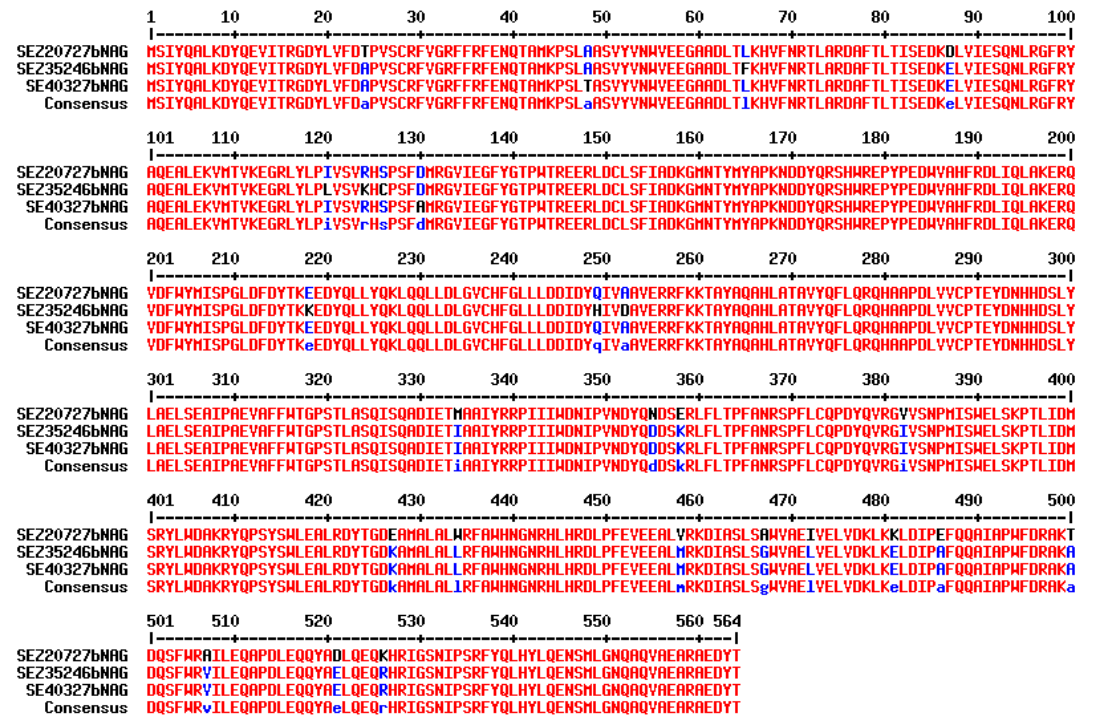


Figure 24 Protein sequences of beta-N-acetylglucosaminidases of SE40327, SEZ20727 and SEZ35246. Sequences were identified using the RAST annotated genomes of the strains and compared using Multalin software (<http://multalin.toulouse.inra.fr/multalin/>). Observed is a number of single amino acid differences between strains throughout the sequences.

Previously it was reported that SER17037 (CECT5772) tested negative for hyaluronidase activity (Speck *et al.*, 2008). Whilst a definitive hyaluronidase gene was not identified from the RAST annotated genome of this strain, a beta-N-acetylglucosaminidase was. A comparison of the amino acid sequence to that of the beta-N-acetylglucosaminidases identified in SE40327 and SEZ20727 revealed identities of 95.21% and 98.23% respectively.

3.3.2. Identification of hyaluronic acid associated genes

The hyaluronan synthase genes for all strains of interest were identified from the sequence of SE4047 *hasA*. The amino acid sequence of SE40327 was identical to the sequences of SEE20561 and SE4047. When compared to SEZ20727 and SER17037 some amino acid substitutions were observed although not at the same positions for each strain as observed in figure 25. Streptococcal HasA contains four predicted active sites (Heldermon, DeAngelis and Weigel, 2001). None of the amino acid substitutions noted between the strains were within any of the sites. By comparing the nucleotide sequences it was determined that the differences in amino acids

between strains is a result of single base substitutions. The remaining hyaluronan related genes *hasB*, *hasD*, *hasE* and both paralogs of *hasC* were also searched for in the sequenced strains using the amino acid sequences from *S. equi* 4047. The genes of *hasB*, *hasD* and *hasE* were identified in SE40327, along with a partial sequence for *hasC1* and the whole of *hasC2*. In SEE20561 *hasB* and *hasD* were identified along with *hasA*, however no sequences were identified as *hasE* or either paralog of *hasC*. All genes were identified in SEZ20727 including both copies of *hasC* and for SER17037 *hasB*, *hasD* and *hasE* were all identified, but neither copy of *hasC*.

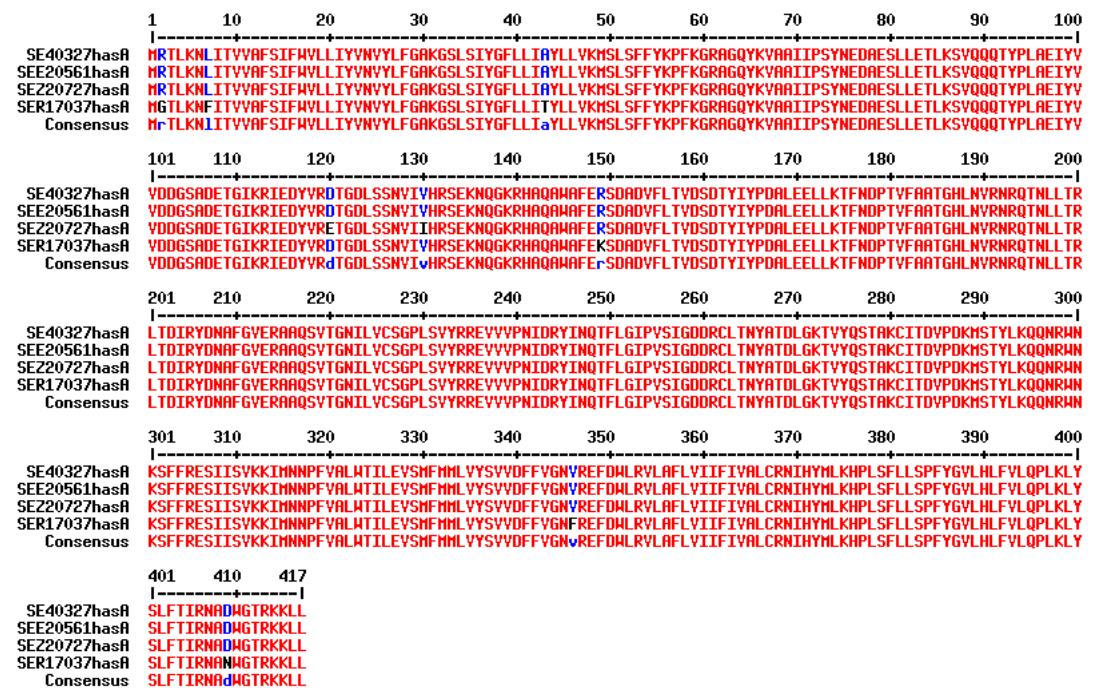


Figure 25 Amino acid sequences of hyaluronan synthase in the genomes of the *S. equi* strains of interest. Sequences were obtained from the RAST annotation of the genome sequences of the strains and were compared using Multalin (<http://multalin.toulouse.inra.fr/multalin/>). Single amino acid differences between the strains are observed throughout the length of the genes.

The organisation of the *has* operons of two subspecies, *S. equi* subsp. *zooepidemicus* ATCC 35246 (SEZ35246) and *S. equi* subsp. *equi* 4047 (SE4047), of which the whole genomes were available were observed. These are presented in figure 26, along with what was available of the operon of the strain of interest SE40327. The same genes were observed upstream in all strains, with a hypothetical protein present immediately upstream of the *hasA* gene which is the first in the operon. All five genes required for HA synthesis were present within the *has* operon in

SEZ35246, whereas in SE4047 the operon appears to consist only of *hasA-C*, with *hasD* and *hasE* observed downstream and on the opposite strand contiguous to the second paralog of *hasC* and a gene for glycerol-3-phosphate dehydrogenase. Due to incomplete sequencing the contig ends following the *hasB* gene in SE40327 and so the remaining operon structure could not be deduced.

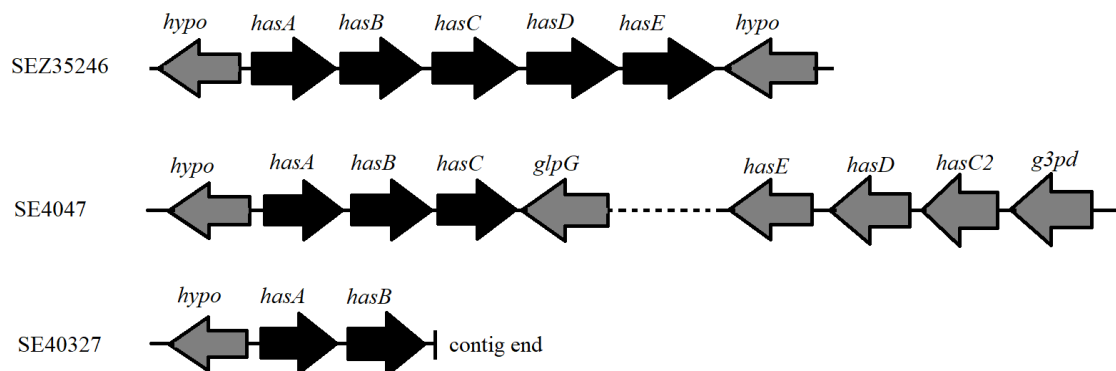


Figure 26 Diagram of the *has* operon of *S. equi* subspecies. The diagram displays the operons from strains of *S. equi* subsp. *zooepidemicus* (SEZ35246), *S. equi* subsp. *equi* (SE4047) and the industrial strain of interest *S. equi* (SE40327). Gene directions are as observed from the RAST annotations of the genomes and are left to right 5' to 3'. The enzymes of the operon are labelled *hasA-E*. Also presented is the preceding gene for a hypothetical protein (*hypo*). The operon in SEZ35246 is complete and contiguous, whereas for SE4047 it is separated into two parts as observed by Blank, Hugenholtz and Nielsen (2008). A GlpG protein (*glpG*) is the first in the series of genes which separate the *has* operon genes. The genes for the paralog of *hasC* (*hasC2*) and a glycerol-3-phosphate dehydrogenase (*g3pd*) are also observed to be neighbouring *hasD* and *hasE*. In SE40327, the contig ends following *hasB*.

The upstream sequences from the start codon site of the *has* operon were analysed in order to identify possible regulatory elements. In SE4047 the ribosome binding site (AGGAGG) was observed 13 bases upstream of the start codon (ATG). Additionally possible sigma factor binding sites -10 and -35 were identified 102 and 129 bases upstream of ATG respectively. In SEZ35246 a possible ribosome binding site (AGG) was identified 16 bases upstream of ATG, whilst the same -10 and -35 sequences were observed at the same upstream positions as SE4047. In SE40327 a possible ribosome binding site (AGG) was observed 21 bases upstream from ATG and the possible -10 (TTGTACCAT) and -35 (TTGCTA) sequences were identified 43 and 65 bases upstream respectively. A diagram of these elements in SE4047 and SE40327 is presented in figure 27.

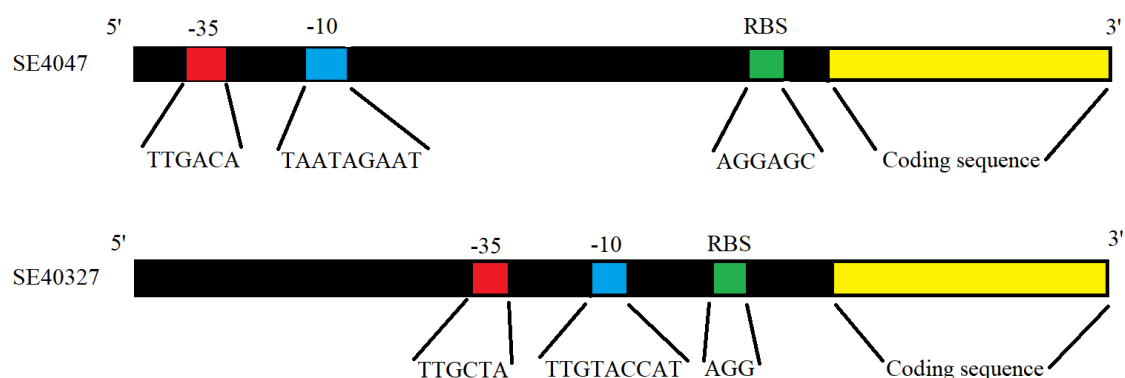


Figure 27 Upstream promoter sequences of the *has* operons of *S. equi* subsp. *equi* strains.

Identified are the possible ribosome binding sequences (RBS) and the -10 and -35 regulatory sequences for strains *S. equi* subsp. *equi* 4047 (SE4047) and the industrial strain of interest *S. equi* 40327 (SE40327). Coding sequence denotes the *hasA* gene and beginning of the *has* operon.

3.3.3. Gene discrepancies between phenotypes

Comparing the RAST annotated gene content of the mucoid variant of SE40327 (40327MV) to the non-mucoid (40327NMV) did not reveal the absence of any genes in 40327NMV which stand out as potential capsule production related candidates for further investigation. Genes with between 0 and 100 percent identity between variants were assessed and 47 genes were identified which were present but did not completely match in sequence identity between the two phenotypes. Some of the differences which were investigated further were found to be due to genes annotated at the end of contigs, meaning part of the sequence in one or both of the phenotypes has been lost during sequencing, assembly and annotation. Those which appeared to be genuine differences occurring in genes not found at the end of contigs were investigated further if considered to be of interest with regards to capsule production. Three genes in particular located adjacent to the *has* operon were investigated. In both strains the first two genes described below were located on the same contig as the *has* operon and not near the ends of the contigs. The third was located in the middle of a contig containing another gene annotated as a hyaluronan synthase but not within the *has* operon. A BLAST of this gene identified it as a glycosyltransferase family 2 of *S. equi* (accession: WP_012679414.1).

3.3.3.1. Collagen-like surface protein

The first selected gene was annotated as a “collagen-like surface protein” in RAST. The sequence was queried in Pfam, Interpro and NCBI Blast to confirm and identified as repeated collagen triple helices and a LPXTG cell wall anchor domain, Gram-positive LPXTG cell wall anchor and a collagen-like cell surface-anchored protein SclH respectively. A comparison of the amino acid sequence (fig 28) of the gene from both phenotypes revealed that in 40327NMV over half of the sequence has been lost (amino acids 1-194). Amino acid sequences for both variants were then compared to the NCBI sequence for LPXTG cell wall anchor domain (WP_012678932.1), which is 414 amino acids long (not shown). The functional domains for this protein are annotated in NCBI as follows. Amino acid range 120-414 is described as DNA polymerase III subunits gamma and tau and 374-411 a gram positive anchor. Accepting these descriptions indicates that in 40327NMV the DNA polymerase region at the N terminus end has been truncated but the gram positive anchor domain is intact in both variants. The LPXTG motif is found at amino acids 382-386 and so has not been lost in either 40327MV or 40327NMV. The amino acid sequence from 40327MV is identical to the sequence from SE4047. Comparing the nucleotide sequences reveals that a total of 576 bases are missing from 40327NMV.

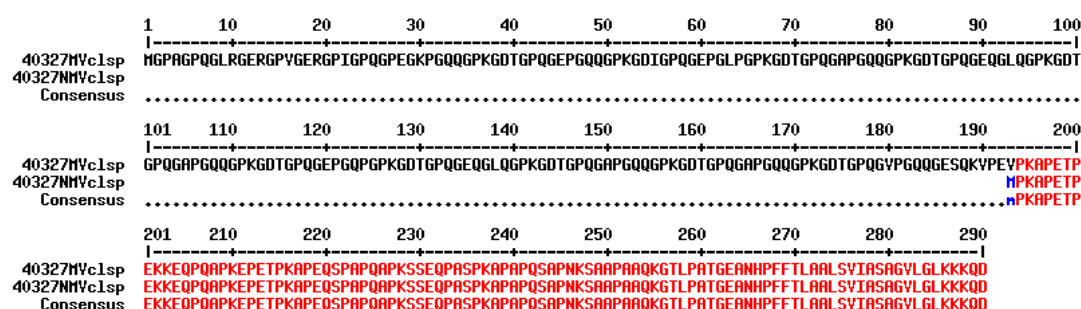


Figure 28 Amino acid sequence of collagen-like surface protein from mucoid (40327MVclsp) and non-mucoid (40327NMVclsp) variants. Sequences were identified from the RAST annotated genomes of the strains and compared using Multalin software (<http://multalin.toulouse.inra.fr/multalin/>). More than half of the protein sequence appears to be missing in the non-mucoid variant, with only the last 97 amino acids at the C terminal present.

The upstream sequences of the genes from both phenotypes were analysed to identify any differences in gene regulation. These were not homologous between the phenotypes. The start codon for this gene has been lost in 40327NMV as has any identifiable ribosome binding site, however -10 and -35 sequences were identified 49 and 78 bases upstream respectively of the beginning of the truncated gene. In 40327MV a possible ribosome binding site (AGCAGG) was identified 17 bases upstream from ATG and -10 and -35 sequences were identified 174 and 196 bases upstream respectively. A diagram of the located sequences in 40327MV is presented in figure 29.

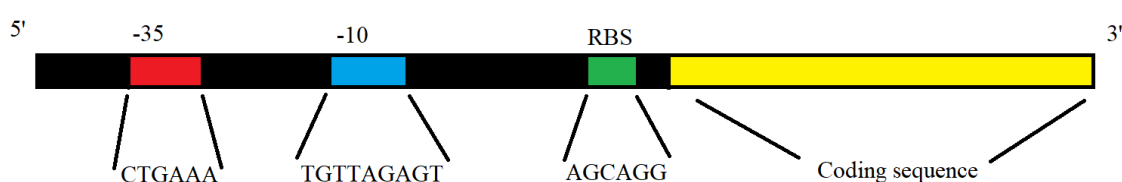


Figure 29 Upstream sequence of collagen-like surface protein of 40327MV. The diagram displays the regulatory sequences of ribosome binding site (RBS) as well as -10 and -35 sequences located upstream of the coding sequence for the collagen-like surface protein gene.

3.3.3.2. M protein trans-acting positive regulator (Mga)

The second gene investigated was annotated in RAST as M protein trans-acting positive regulator (Mga). This annotation was confirmed using NCBI, Interpro and Pfam and genes were also located in the middle of contigs and upstream from *hasA* in both phenotypes. The sole amino acid sequence difference between mucoid and non-mucoid phenotypes is a substitution observed at position 17, with 40327MV possessing a leucine at this position and 40327NMV a tryptophan. The same gene in SE4047 also possesses a leucine at this position of the protein. The region 8-64 of Streptococcal Mga-like regulatory protein (WP_012678912.1) is designated a DNA binding helix-turn-helix (HTH) domain in NCBI and the substitution of leucine for tryptophan introduces two aromatic rings in place of methyl groups into this sequence region. Two other genes of 40327MV were annotated as Mga by RAST, however these were not confirmed when queried in NCBI. Comparison of the nucleotide sequences reveals that the amino acid change is a result of a single nucleotide

substitution from a thymine to a guanine at position 50 (data not shown). The Mga operon and regulon has been extensively studied in *S. pyogenes*, but little has been investigated with regards to this transcriptional regulator in Group C streptococci. A diagram of the operon identified in SE40327 is presented in figure 30. In both species *mga* is the first gene in the operon and is immediately followed by the gene for antiphagocytic M protein (*emm*) or M-like protein (*sem*).

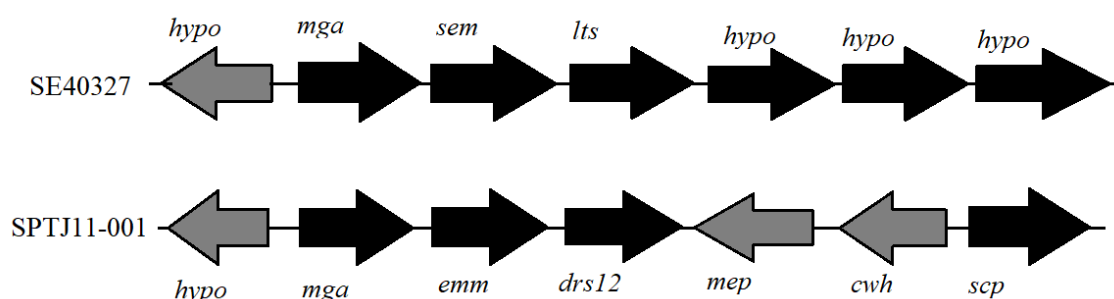


Figure 30 The *mga* operons of *S. equi* subsp. *equi* 40327 (SE40327) and *S. pyogenes* TJ11-001 (SPTJ11-001). A hypothetical protein (*hypo*) is observed in both strains prior to the start of the operon in which *mga* is the first gene. In both cases *mga* is immediately followed by the M proteins of the species (*sem/emm*). However while C5a peptidase (*scp*) is included in SPTJ11-011 the operon of SE40327 appears to only include a leucine tRNA synthase (*lts*) and more hypothetical proteins. In SPTJ11-011 *drs12*, *mep* and *cwh* denotes streptococcal competence protein, mobile element protein and phage cell wall hydrolase respectively.

3.3.3.3. RNA polymerase sigma-70 factor

A third gene identified as being different between phenotypes was an RNA polymerase sigma-70 factor. This was confirmed using pfam and Interpro as described previously and matched 100% to the NCBI accession WP_012679409.1. A comparison of the amino acid sequences between the phenotypes revealed that the 40327NMV sequence was missing the first 11 amino acids and the 12th possessed a leucine in place of a methionine. In both phenotypes the gene immediately upstream within the same contig from the sigma factor is a pullanase. These were also compared between phenotypes and were found to be identical, comparing the nucleotide sequences of both genes in tandem between phenotypes revealed a gap of 17 nucleotides in the non-mucoid phenotype.

3.3.4. The *hasC* genes

Intact sequences for both paralogs of the *hasC* genes were not able to be identified in SE40327 or any other of our strains of interest apart from SEZ20727 so the presence of two copies in these strains have not been confirmed. The complete genomes of *S. equi* subsp. *equi* ATCC 39506 (SEE39506), *S. equi* subsp. *equi* 4047 (SE4047), *S. equi* subsp. *zooepidemicus* ATCC 35246 (SEZ35246) and *S. equi* subsp. *zooepidemicus* MGCS10565 (SEZMCGS) are available on NCBI. These were uploaded to RAST for annotation as reference strains. A comparison of the amino acid sequences of the paralogs of SE4047 shows variation in sequence between amino acids 291-308. The *hasC2* paralog is also missing the last two amino acids upon alignment of the sequences. In addition to this at position 11 an alanine in HasC1 is replaced with a threonine in HasC2 (fig 31), a substitution which introduces a hydroxyl group and a methyl group. A comparison of the nucleotide sequences reveals that this substitution is the result of a single nucleotide change at position 31 where a guanine in *hasC1* is substituted for adenine in *hasC2*. The sequences are then identical until nucleotide position 870 where there is variation until the end of the gene sequence (909) and *hasC2* is missing the final 5 nucleotides present in *hasC1*. A comparison of the SE4047 paralog nucleotide sequences with their corresponding genes in SEE39506 reveals complete identity of the genes between these strains, including the region of variation at the C-terminus end. The amino acid substitution observed between the SE4047 paralogs is within a reported active site of HasC.

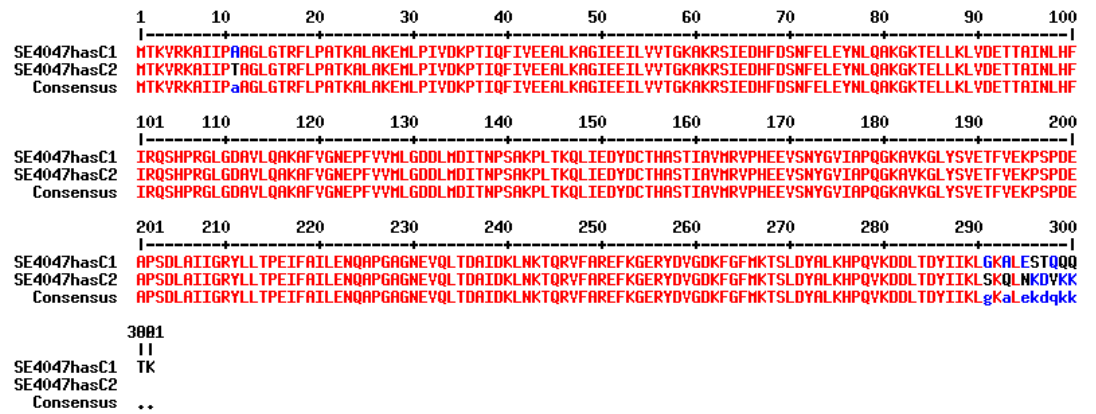


Figure 31 Amino acid sequence of the HasC paralogs in SE4047. The sequences were identified from the RAST annotated genome of SE4047 and compared using Multalin software (<http://multalin.toulouse.inra.fr/multalin/>). Variation in amino acid sequence is observed at the C terminus from amino acid 291. Additionally, there is a difference in amino acid at position 11.

The paralog sequences of SEZMGCS are identical until amino acid 291 where differences are observed. Comparing the nucleotide sequences reveals a number of single and double base substitutions throughout, however none of these appear to result in an amino acid change until nucleotide 870 where the sequences possess greater variation. Similar is observed for the paralogs of SEZ35246, however only one nucleotide substitution is noted at position 861 where a thymine is substituted for cytosine in *hasC2* prior to the variation in sequences beginning at nucleotide 870 as seen in other strains. Meanwhile SEZ20727 has around 5 base substitutions between paralogs prior to nucleotide number 870 in comparison to 13 substitutions in SEZMGCS. Comparing the HasC1 amino acid sequences of all three *S. equi* subsp. *zooepidemicus* strains reveals only two amino acid differences at positions 223 and 284. At 223 SEZ35246 possesses an asparagine whereas SEZ20727 and SEZMGCS possess lysine and at 284 the strains possess alanine or threonine respectively. The amino acid sequences of SEZ20727 and SEZMGCS are identical. However by comparing the nucleotide sequences of these strains it becomes clear that differences exist which may not affect the amino acid sequence. A comparison of the *hasC1* sequences of SEZ20727 and SEZ35246 revealed 15 nucleotide differences. Whereas SEZ20727 and SEZMGCS have 11 differences between them and SEZMGCS has 14 nucleotides different from SEZ35246. Similar can be said for the *hasC2* paralog sequences of these strains with 16 differences observed between

SEZ20727 and SEZ35246, 6 differences between SEZ20727 and SEZMGCS and 22 between SEZ35246 and SEZMGCS. In both paralogs it appears that SEZ20727 is closer in sequence to SEZMGCS than SEZ35246. However, although numerous differences in nucleotide sequences are observed between strains, the amino acid sequences are almost identical with only the previously mentioned amino acid differences observed.

3.3.5. The MurA paralogs

The sequences of the MurA paralogs in all the strains were identified from the RAST annotated genomes. Each paralog for the *S. equi* strains of interest (SE40327, SEE20561, SEZ20727 and SER17037) were compared at protein and nucleotide levels along with the sequences from *S. equi* subsp. *equi* 4047 (SEE4047) and *S. equi* subsp. *zooepidemicus* ATCC 35246 (SEZ35246) which were used as reference strains as the whole genome sequences are available for both.

3.3.5.1. MurA1

The amino acid sequence of MurA1 in the strains is presented in Fig 32. The sequences of SE40327, SEE20561 and SEE4047 were identical to each other, as were SEZ20727 and SEZ35246. The sequence of SER17037 possessed 99.52 % identity to the *S. equi* subsp. *equi* strains and 99.05 % identity to *S. equi* subsp. *zooepidemicus* strains. A total of 6 differences in amino acid sequences were observed between the *equi* and *zooepidemicus* strains. SER17037 matched the *equi* strains sequences in four of these instances (positions 163, 247, 286 and 360) but matched *zooepidemicus* strains in the other two (positions 80 and 318). The MurA1 sequence of SE40327 identified 100% with an *S. equi* sequence for UDP-N-acetylglucosamine 1-carboxyvinyltransferase (reference WP_012679232.1). According to the NCBI record for this sequence the active sites of the enzyme are found at amino acid positions 23, 92, 96, 121, 126, 162, 165, 166 and 306. None of the amino acid substitutions observed between *S. equi* subsp. *equi* and *S. equi* subsp. *zooepidemicus* occur at these positions. Additionally,

none of these substitutions occur at positions designated ‘hinge sites’.

A comparison of the nucleotide sequences of this gene between strains reveals the amino acid changes to be the result of single base differences throughout the gene.

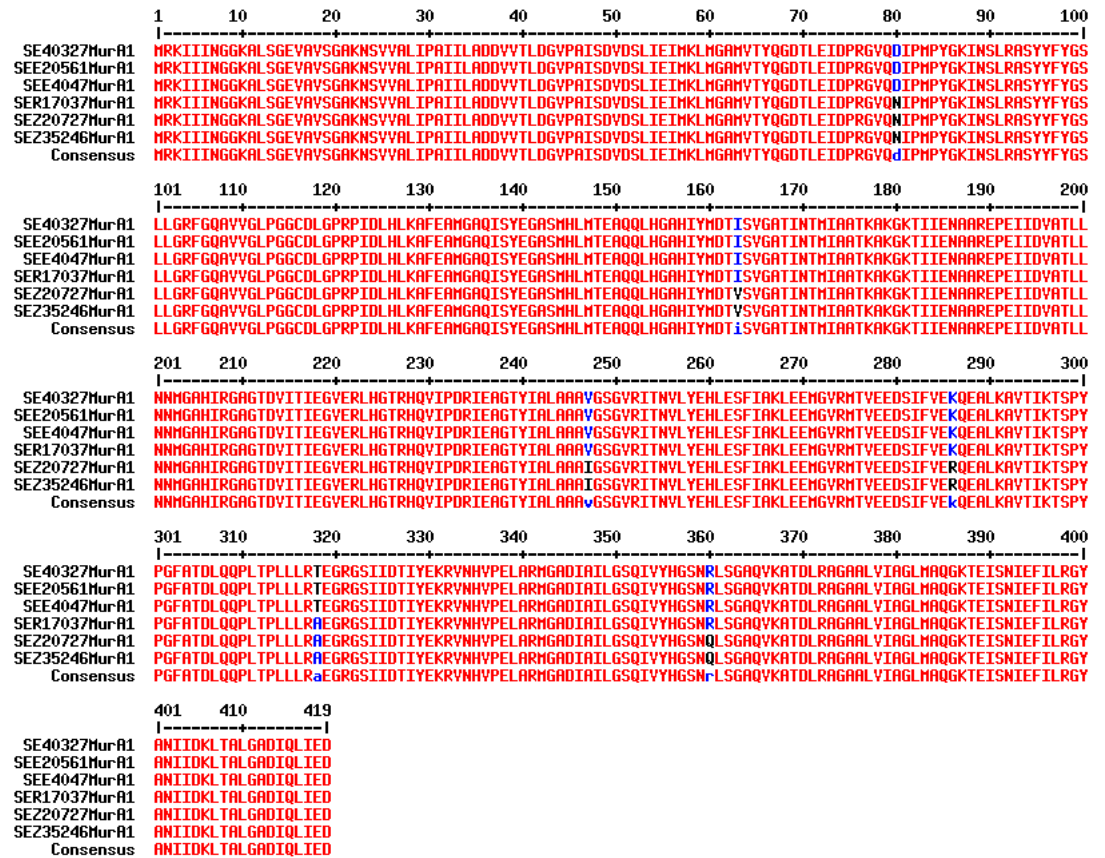


Figure 32 Amino acid sequences of MurA1 of *S. equi* strains. Sequences were identified from RAST annotations of the genomes and compared using Multalin software (<http://multalin.toulouse.inra.fr/multalin/>). Amino acid substitutions are observed between *S. equi* subsp. *equi* strains compared to *S. equi* subsp. *zooepidemicus* strains. *S. equi* subsp. *ruminatorum* matches the *equi* strains at four of these positions and *zooepidemicus* strains at two.

A phylogenetic tree was constructed using the *murA1* gene sequences of the strains mentioned above and *S. equi* subsp. *zooepidemicus* MGCS10565 in order to visualise the relationship between strains (fig 33). From this a clear distinction between *S. equi* subsp. *equi* species and *S. equi* subsp. *zooepidemicus* species is observed and *S. equi* subsp. *ruminatorum* possesses greater relation to the *equi* species.

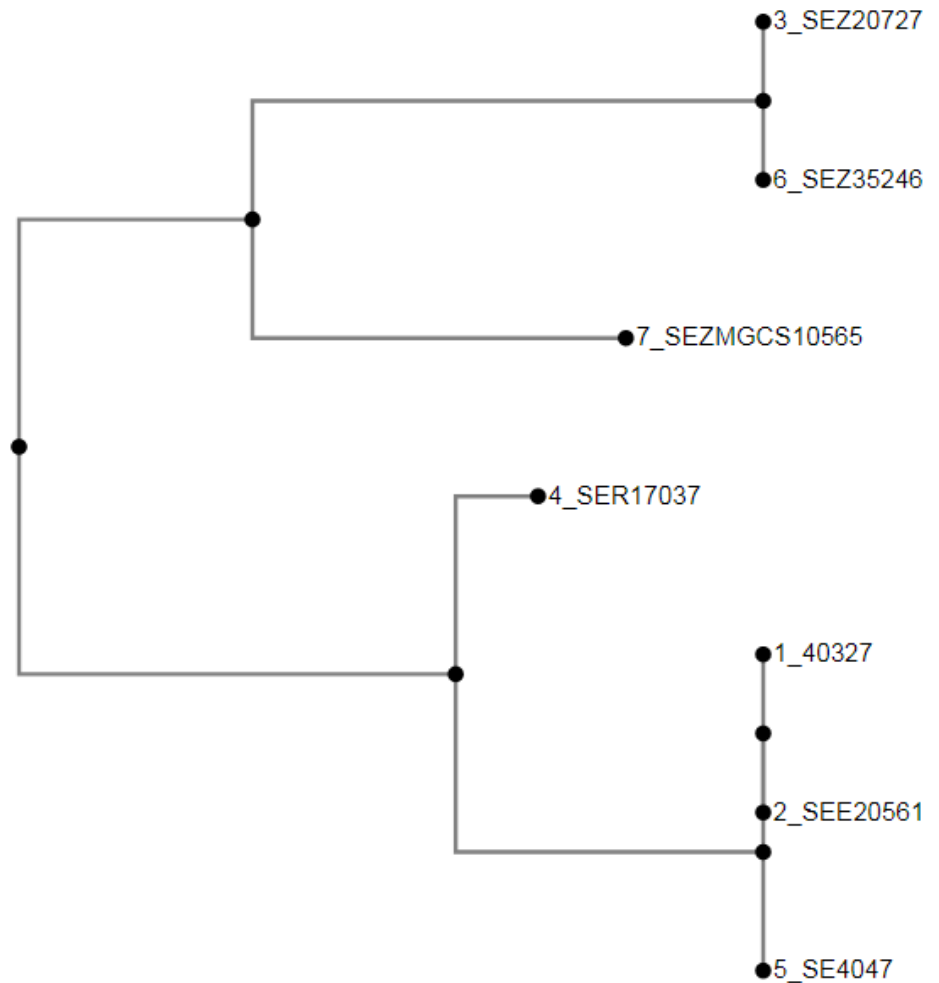


Figure 33 Phylogenetic tree constructed using the *murA1* nucleotide sequences from the *S. equi* strains. Sequences were aligned and trees were constructed using the MAFFT online alignment software (<https://mafft.cbrc.jp/alignment/server/>) and viewed using Phylo.io. A clear differentiation is observed between *S. equi* subsp. *equi* strains (40327, SEE20561 and SE4047) and *S. equi* subsp. *zooepidemicus* strains (SEZ20727, SEZ35246 and SEZMGCS10565). The *murA1* gene of *S. equi* subsp. *rumintorum* (SER17037) appears to be more closely related to the *equi* strains than the *zooepidemicus*.

3.3.5.2. MurA2

A BLAST search of the MurA2 amino acid sequence from SE40327 resulted in a maximum identity of 99.76% with *S. equi* UDP-N-acetylglucosamine 1-carboxyvinyltransferase (WP_012679395.1). A comparison of the amino acid sequences of the *S. equi* species along with reference strains described above and the NCBI reference sequence is presented in fig 34. From this it is observed that a single amino acid difference is present between SE40327 and

WP_012679395.1 at position 180. Additionally, when all are aligned it is observed that the sequence for SEZ20727 possesses an extra four amino acids at the N terminus and a valine in place of a methionine at what is designated position one for all the other strains. The active sites for this sequence according to the NCBI record are at positions 24, 96, 100, 125, 126, 127, 128, 129, 130, 164, 167, 168 and 309 and so the amino acid changes noted are not within an active region of the protein nor are they located at positions designated 'hinge sites'. The substitution described at position 180 appears to only be present in SE40327 where the amino acid present is a proline as opposed to an alanine. An additional 5 variations in amino acids between the sequences are present at positions 10, 57, 153, 271 and 294. Once again SER17037 appears to differ slightly from both *S. equi* subsp. *equi* and *S. equi* subsp. *zooepidemicus* species at different amino acid positions. A comparison of the nucleotide sequences showed SEZ20727 to possess 12 extra nucleotides at the beginning of the sequence. A series of single and double base substitutions are observed throughout the genes, some of which are possibly the cause of the amino acid differences observed between the protein sequences.

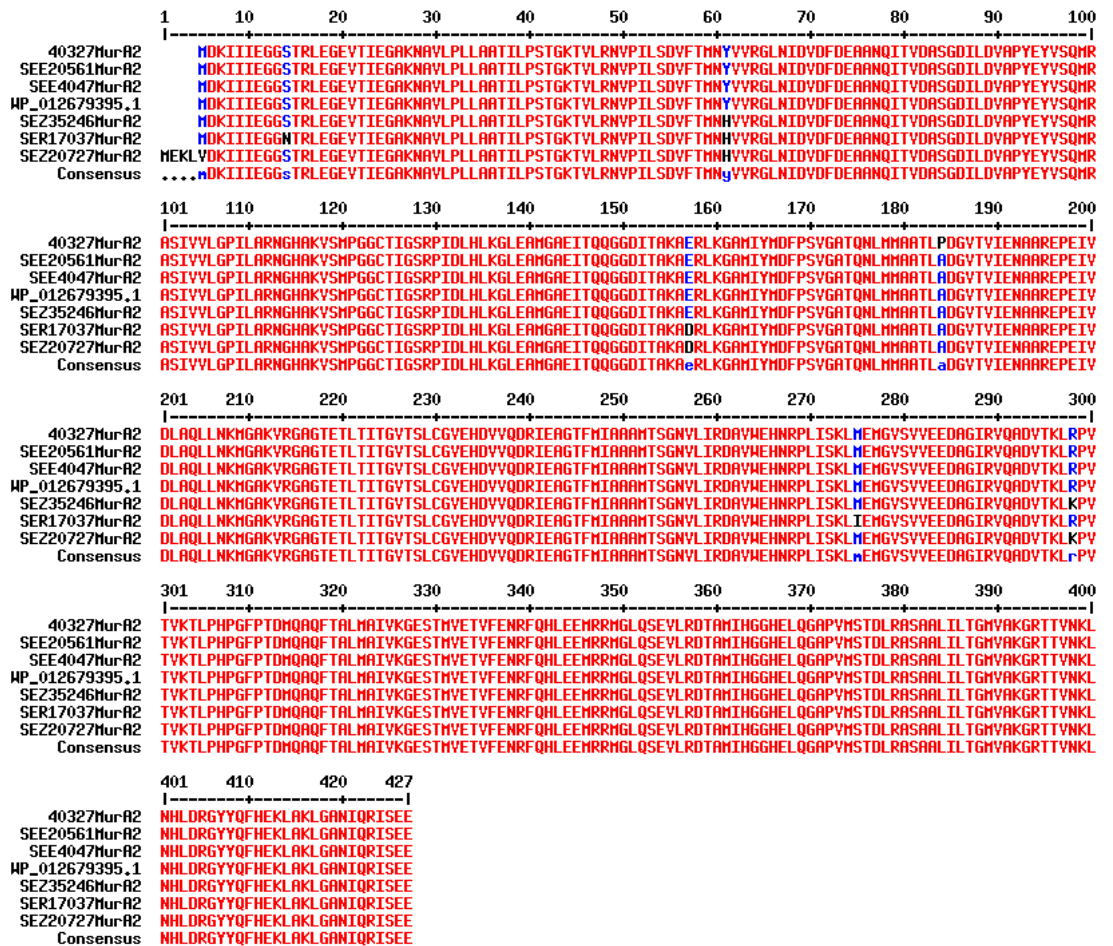


Figure 34 Amino acid sequences of MurA2 of *S. equi* strains and NCBI sequence

WP_012679395.1. Sequences were identified from RAST annotations of the genomes or from NCBI and compared using Multalin software (<http://multalin.toulouse.inra.fr/multalin/>). SEZ20727 is observed to possess four extra amino acids compared to the other strains as well as a valine in place of a methionine at what corresponds to position 1 of the protein in the other strains. A series of amino acid substitutions between some of the strains are present throughout the sequences.

A phylogenetic tree was constructed using nucleotide sequences as described previously in order to visualise the relationship between strains based on *murA2* (fig 35). From the tree it is noted that SER17037 diverges from the other *S. equi* strains. *S. equi* subsp. *equi* sequences are closely clustered whereas – in contrast to MurA1 – the separation between the *equi* and *zooepidemicus* is not evident. Instead, MurA2 appears to be undergoing gradual divergence within *S. equi* subsp. *zooepidemicus*.

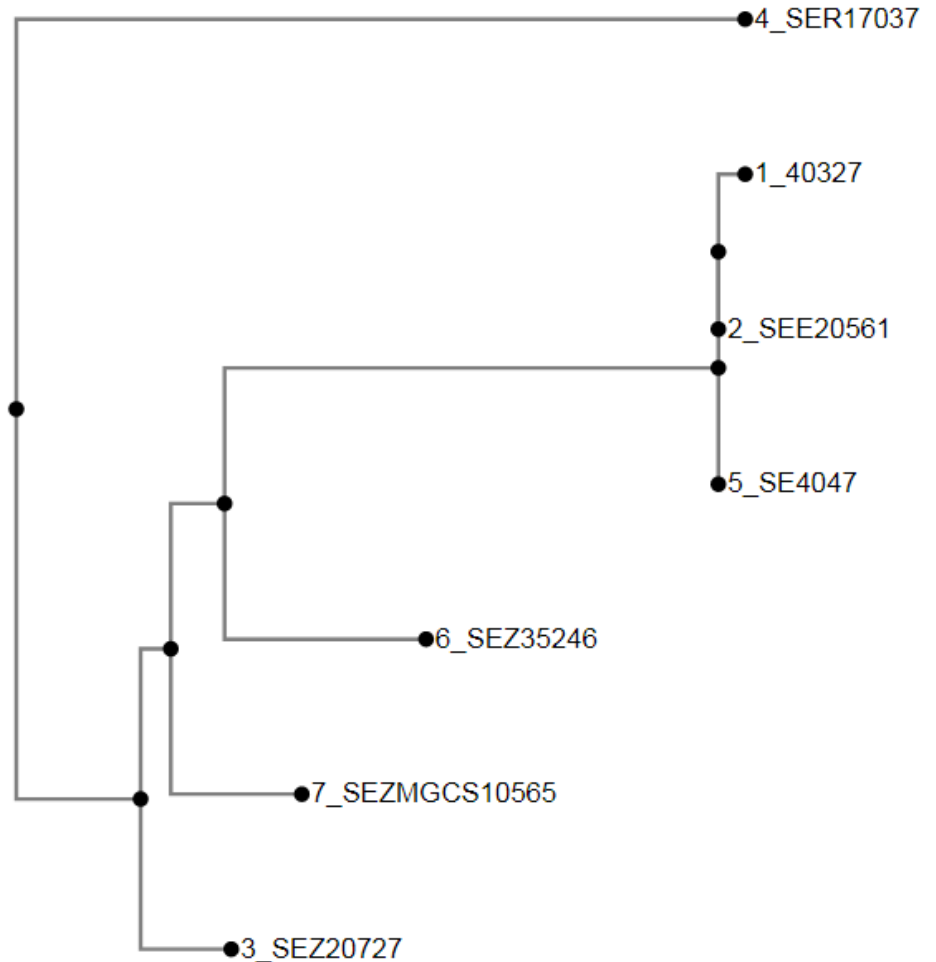


Figure 35 Phylogenetic tree constructed using the *murA2* nucleotide sequences from the *S. equi* strains. Sequences were aligned and trees were constructed using the MAFFT online alignment software (<https://mafft.cbrc.jp/alignment/server/>) and viewed using Phylo.io. *S. equi* subsp. *ruminantium* (SER17037) appears to be separate from the other *S. equi* strains. The sequences of strains of *S. equi* subsp. *equi* (40327, SEE20561 and SE4047) are grouped together as seen in the tree for *murAI*, however they also appear to be more closely related to the *S. equi* subsp. *zooepidemicus* strains than the *murAI* sequences.

3.4. Discussion

3.4.1. Hyaluronidases of *Streptococci*

‘Hyaluronidase’ was previously used as a general term to describe enzymes which degrade HA, though their activity is often not entirely specific to HA alone (Hynes and Walton, 2000). Thus hyaluronidases are a family of enzymes which have been established to comprise of three main groups, one of which is designated the bacterial hyaluronidases or hyaluronate lyases (Menzel and Farr, 1998). Hyaluronidases are speculated to

contribute to the pathogenicity of bacteria through degradation of HA within host tissues and as such are considered virulence factors (Starr and Engleberg, 2006). In bacteria there are different types of hyaluronidases, which are either chromosome-encoded extracellular enzymes thought to have roles in host invasion or enzymes encoded by integrated phage DNA in the bacterial genome (Hynes, Hancock and Ferretti, 1995).

The phages of hyaluronic acid producing bacteria are thought to encode hyaluronidase enzymes in order to facilitate access to the cell through the capsule (Mylvaganam *et al.*, 2000). Hyaluronidases of both bacterial and phage origin have been found to vary greatly in sequence length and molecular weight (Girish and Kemparaju, 2007) and sequence comparisons have found the bacterial hyaluronidases to diverge substantially in some cases, whereas bacteriophage hyaluronidases possess a high degree of homology (Hynes and Walton, 2000). Despite claims that several bacterial hyaluronidases are well characterised, the use of different names to identify the same enzymes results in confusion around whether an enzyme is a phage hyaluronidase or not and the significance of the differences between them is left unclear.

Similar enzymes between strains are labelled as hyaluronidases, phage hyaluronidases, hyaluronate lyases and hyaluronoglucosaminidases, and the same sequence is often characterised as more than one of these depending on the database used to search and compare. This results in difficulty in knowing whether enzymes identified in newly sequenced genomes may be true bacterial hyaluronidase genes or hyaluronidases from bacteriophage DNA. Fig 36 shows the amino acid sequences of all annotated hyaluronidase genes in *S. pyogenes* strain SSI-1. The genes are annotated in NCBI as either putative hyaluronidase, putative hyaluronidase (phage associated) or putative hyaluronoglucosaminidase. A comparison of the sequences indicates areas of homology between all but two of the genes, those annotated hyaluronidase (SSI-1_hyal) and putative hyaluronidase (phage associated) (SSI-1_putphagehyal3).

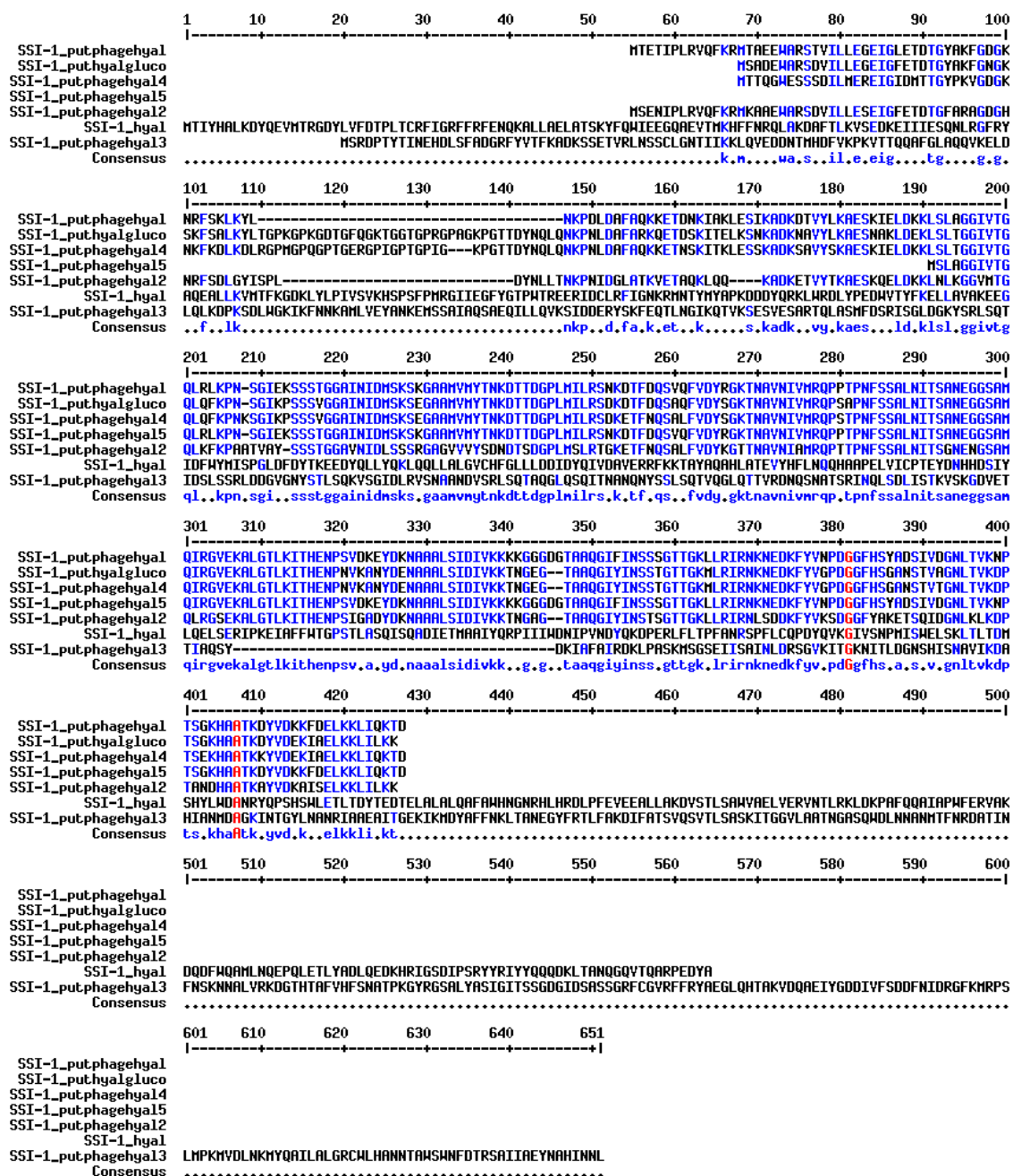


Figure 36 Amino acid sequences of hyaluronidase genes of *S. pyogenes* strain SSI-1. All sequences were identified from NCBI where they were annotated as either putative hyaluronidase (hyal), putative hyaluronidase (phage associated) (putphagehyal) or putative hyaluronoglucosaminidase (puthyalgluco). Areas of sequence identity are present between most of the sequences apart from SSI-1_putphagehyal3 and SSI-1_hyal which are longer and do not possess areas of homology to either each other or the other sequences. Sequences were aligned using Multalin software (<http://multalin.toulouse.inra.fr/multalin/>).

Having identified a number of genes annotated as hyaluronidases in the SE40327 genome by RAST it was necessary to be confident that investigations involved ‘true’ enzymes going forward. Thus the use of three separate protein databases to confirm the identification of the genes

provided a means of ensuring certainty of the original annotations. While not full proof, this method uses current knowledge of the hyaluronidase families as a first step towards confirmation and resulted in identification of two genes to concentrate initial analyses.

3.4.1.1. SE40327hyal1 and SE40327hyal2, true hyaluronidases?

The two accepted hyaluronidase sequences of SE40327 were compared to the sequences isolated from SSI-1. The greatest match for SE40327hyal1 was to putative hyaluronidase (phage associated) (NCBI: SPs0648) which matched 100% of the query cover at an identity of 75.47%. Upon comparing SE40327hyal1 against genes of *S. equi* subsp. *equi* ATCC 39506 it was found to match 88.14% to 100% of the sequence for the gene SE071780_01756 which is annotated in NCBI as a hyaluronoglucosaminidase. A comparison of the amino acid sequences of SE40327hyal1, Sps0648 and SE071780_01756 is presented in fig 37. The majority of the differences observed between the sequences appear to be towards the N terminal of the proteins and SPs0648 is missing a section of approximately 30 amino acids.

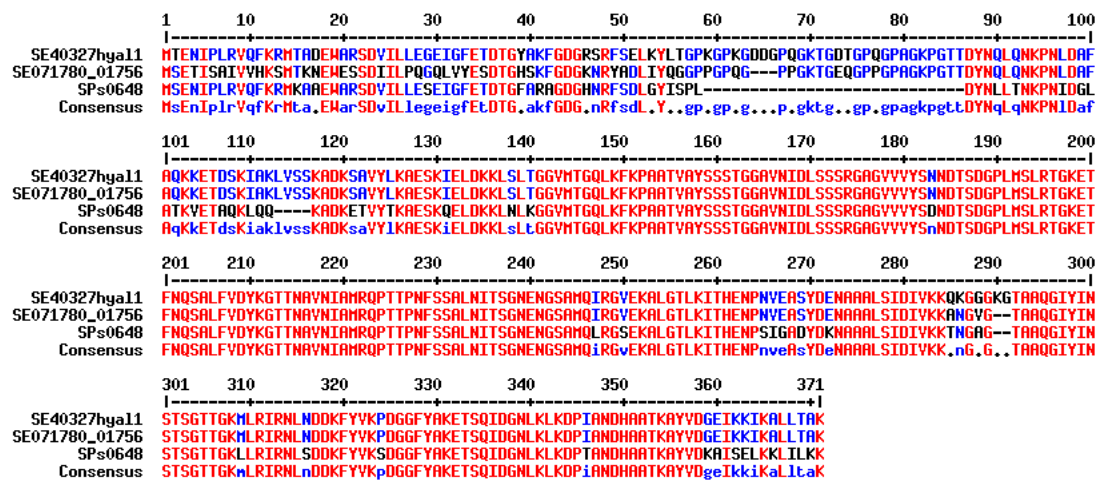


Figure 37 Amino acid sequences of hyaluronidase genes of SE40327 (SE40327hyal1), *S. equi* subsp. *equi* ATCC 39506 (SE071780_01756) and *S. pyogenes* strain SSI-1 (SPs0648). SE40327hyal1 was identified using the RAST annotated genome of SE40327 and the sequences of the other strains were identified from the nucleotide sequences in NCBI (<https://www.ncbi.nlm.nih.gov/>). Sequences were aligned using Multalin software (<http://multalin.toulouse.inra.fr/multalin/>).

The second SE40327 hyaluronidase sequence (SE40327hyal2) matched 83.75% to 95% query cover of a separate putative hyaluronidase (phage associated) of SSI-1 (SPs0927). Whereas when compared to the sequences of *S. equi* subsp. *equi* ATCC 39506 it matched 100% of the whole query to the gene SE071780_02165 which is annotated as a hyaluronoglucosaminidase. The amino acid sequences are compared in fig 38. The *S. equi* strains appear to have an extra 14 amino acids at the N-terminus, however aside from this observation there appears to be minimal discrepancy between the *S. equi* and *S. pyogenes* strains.

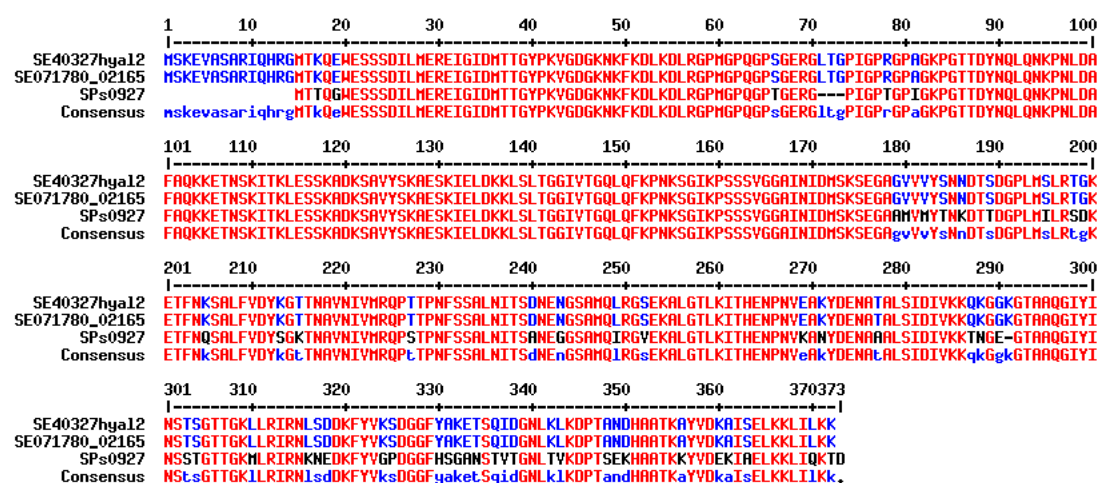


Figure 38 Amino acid sequences of genes from SE40327 (SE40327hyal2), *S. equi* subsp. *equi* ATCC 39506 (SE071780_02165) and *S. pyogenes* SSI-1 (SPs0927). SE40327hyal2 was identified using the RAST annotated genome of SE40327 and the sequences of the other genes were identified from the nucleotide sequences in NCBI (<https://www.ncbi.nlm.nih.gov/>). Sequences were aligned using Multalin software (<http://multalin.toulouse.inra.fr/multalin/>).

Previously it has been claimed that *S. pyogenes* possesses two hyaluronidase genes, one bacterial, designated the name *hylA* and one which is encoded by phage DNA (Hynes *et al.*, 2000). However there are seven genes annotated as hyaluronidases or hyaluronoglucosaminidases in the NCBI entry for *S. pyogenes* strain SSI-1. This is another example of either the vast differences between enzymes within strains of bacteria or of the lack of confirmatory information and consistency when identifying hyaluronidase enzymes. As previously mentioned the many names associated with hyaluronidases of both bacterial and phage origin leads to confusion

surrounding what is a true hyaluronidase enzyme and what is not. The annotation of most hyaluronidase genes appears still to be mainly speculative and much more is yet to be unveiled about the variety and function of these enzymes in different bacteria.

Previously it was stated that the molecular weight of bacteriophage hyaluronidases is generally 36-40 kDa, with chromosomal bacterial hyaluronidases averaging above this (Hynes, 2000). Both SE40327hyal1 and SE40327hyal2 are approximately 40 kDa in weight, indicating that they are consistent with phage-associated enzymes. The first Streptococcal hyaluronidase enzyme to be sequenced was L20670 from *S. pneumoniae* which is 949 amino acids long and possesses a molecular weight of above 100 kDa. SE40327hyal1 and SE40327hyal2 possess less than 35% homology to this sequence, further suggesting their origins are phage rather than bacterial. Whether these enzymes possess any functionality is unknown and experimental data would be required. In addition, the other enzymes annotated by RAST which did not meet the criteria for further investigation may be being overlooked due to ambiguity around the general annotation of bacterial hyaluronidases. However the two confirmed sequences provide a place to start in the evaluation of hyaluronidase activity in this strain of *S. equi*. No confirmed hyaluronidase genes were identified in either SEZ20727 or SER17037 to compare the SE40327 genes to. However two genes were identified in SEE20561 by the RAST annotation and confirmed as described previously. These were labelled SEE20561hyal1 and SEE20561hyal2. At amino acid level SE40327hyal1 and SEE20561hyal1 possessed 67.39% identity and as observed between SE40327 and *S. equi* subsp. *equi* ATCC 39506 the greatest variation between sequences was observed at the N terminal end of the protein. Whereas the hyal2 genes of SE40327 and SEE20561 were completely identical. No data on the active sites for these enzymes are available and so it is not possible to speculate whether the areas of variation between *S. equi* strains are significant with regards to enzyme function. However now that these phage hyaluronidase genes have

been identified within the *S. equi* strain of interest and appears to also be present in other strains this provides a platform for moving forward with establishing functionality and possible effects on HA capsule production in this strain.

3.4.1.2. Beta-N-acetylglucosaminidases

While hyaluronidase enzymes were not identified for SEZ20727 and SER17037 by RAST, all strains of interest (SE40327, SEE20561, SEZ20727 and SER17037) were found to possess an enzyme matching to β -N-acetylglucosaminidase (bNAG). These enzymes are another member of the family of glycoside hydrolases (EC 3.2.1) which includes hyaluronidases (Zhang *et al.*, 2018). A comparison of the amino acid sequences of these genes is presented in fig 39. A high degree of homology is apparent with only 28 amino acids which vary between strains observed of 564. The sequences of SE40327 and SEE20561 are identical and as seen in other protein sequences SER17037 differs from both *S. equi* subsp. *equi* and *S. equi* subsp. *zooepidemicus* strains at different amino acid positions.

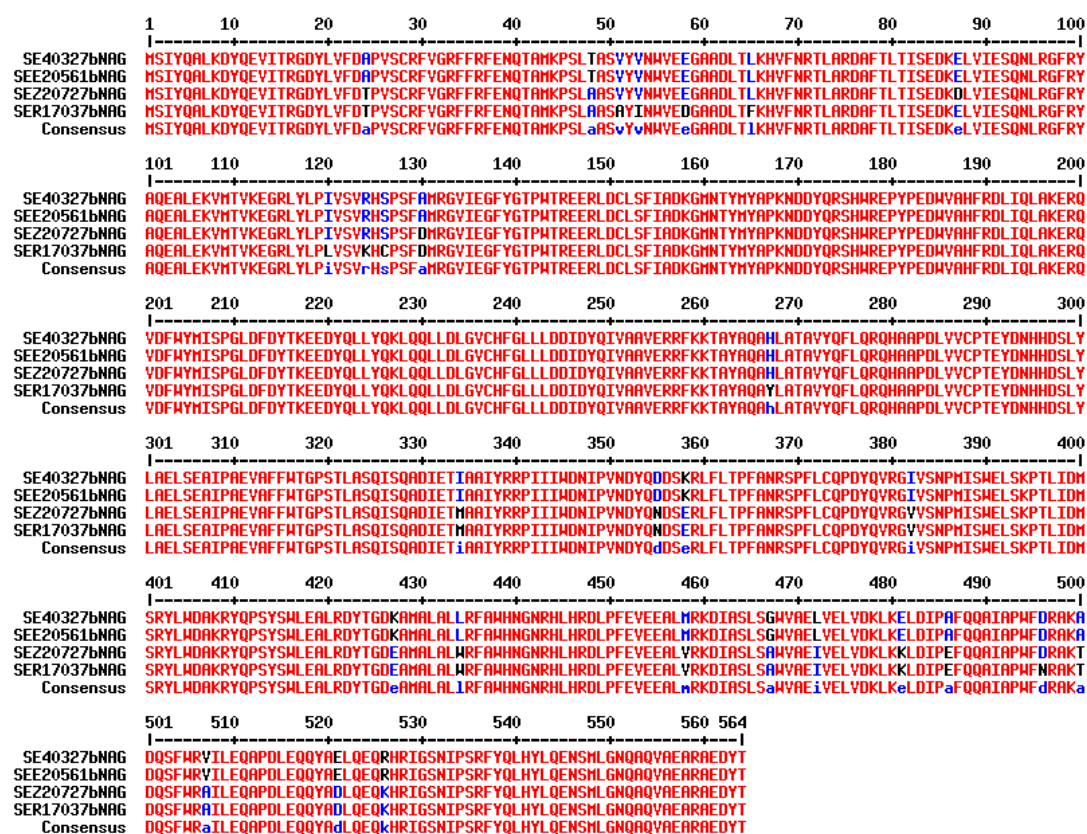


Figure 39 Amino acid sequences of genes annotated as β -N-acetylglucosaminidases in strains SE40327 (SE40327bNAG), SEE20561 (SEE20561bNAG), SEZ20727 (SEZ20727bNAG) and SER17037 (SER17037bNAG). Sequences were identified from the RAST annotations of the genomes and compared using Multalin software (<http://multalin.toulouse.inra.fr/multalin/>). A high degree of homology between the strains is observed and *S. equi* subsp. *equi* strains are completely identical.

Previously it was observed that SER17037 did not exhibit hyaluronidase activity but both SEE20561 and SEZ20727 did (Speck *et al.*, 2008). In the case of SER17037 this claim is confirmed by a lack of hyaluronidase gene identification. It is possible that the lack of hyaluronidase genes identified in SEZ20727 is due to the sequencing method meaning that some genes which fall between or at the end of contigs may be lost, as well as those which possess multiple repeats within the genome. Whilst there has been speculation that bNAGs possess the ability to degrade hyaluronan, the lack of hyaluronidase activity observed in SER17037 would indicate that this is not the case for this enzyme. This conclusion is supported by previous observations in *Streptococcus* species where bNAGS were found to have no effect on hyaluronan (Sheldon *et al.*, 2006; Sun *et al.*, 2015). Due to this, whilst the apparent conservation of this

enzyme between subspecies of *S. equi* is interesting, it is outwith the scope of this project and was not investigated further.

3.4.2. Identification of hyaluronic acid associated genes and the *has* operon

HA production in *S. equi* species involves the following five enzymes, hyaluronan synthase (*hasA*), UDP-glucose dehydrogenase (*hasB*), UDP-glucose pyrophosphorylase (*hasC/galU*), (*hasD/glmU*) and phosphoglucosyltransferase (*hasE/pgi*) following the biosynthetic pathway presented in Figure 40 (Chapter 4). Hyaluronan synthase is responsible for assembling the HA chain and the other enzymes are responsible for producing the two precursors (Chen *et al.*, 2009b). Previously it has been described that the *has* operon of *S. equi* subspecies differs between *zooepidemicus* and *equi* in that all five genes are encoded within a single contiguous operon in *zooepidemicus* but in *equi* the operon contains only *hasA*, *hasB* and *hasC*, with *hasD* and *hasE* being organised within a separate operon (Blank, Hugenholtz and Nielsen, 2008). Unfortunately in all of the genes of interest the beginning of the *has* operon was at the end of contigs and so the structure of the operon could not be confirmed as previously described. Due to the fragmentation of the DNA during paired-end sequencing, genes may be lost and operons may be split between contigs which prevents the analyses of operon structure (Alkan, Sajjadian and Eichler, 2011; Denton *et al.*, 2014). Completion of the genome would require further rounds of sequencing which unfortunately the budget for this project would not allow.

3.4.3. Gene discrepancies between phenotypes

Whilst no obvious discrepancies were observed in the genomes of the two phenotypes of SE40327 in terms of genes involved directly in HA production a number of variations were observed in other genes. Deciding which genes were of interest and confirming the variation between phenotypes was challenging, however three genes were eventually identified and investigated further.

3.4.3.1. Collagen-like surface protein

This gene was selected for investigation initially due to the fact it is located on the same contig and upstream of the *has* operon. Following further analysis it was discovered that the sequence matched to an LPXTG motif. LPXTG motifs are commonly found in surface proteins of Streptococci and are recognised by sortases during the covalent attachment of complete surface proteins such as virulence factors like M protein to the cell wall (Scott and Barnett, 2006). It has previously been observed that in mutants deficient in the LPXTG motif surface proteins fail to associate with the cell wall, highlighting its importance in the sorting and arrangement of surface proteins (Novick, 2000). The nucleotide sequence of the gene from 40327MV matches 100% to the collagen-like cell surface-anchored protein sclH of *S. equi* subsp. *equi* (accession: VEH29797.1). The lack of identifiable start codon or ribosome binding site in 40327NMV suggests that transcription of this gene may have been lost in this phenotype. Another enzyme involved in the placement of surface proteins is foldase which is discussed in chapter 5.

3.4.3.2. M protein trans-acting positive regulator Mga

Mga is a regulator of multiple genes that functions to activate the transcription of virulence factors in response to environmental stimuli (McIver, Thurman and Scott, 1999). The introduction of aromatic rings into a protein by amino acid substitution is likely to promote changes to the structure, stability, flexibility and hydrophobicity. As the substitution appears to have occurred within the DNA binding helix-turn-helix region it is possible this could impede DNA binding and regulation of transcription. “Phase-locked” variants of *Streptococcus* with respect to M protein and HA capsule were often observed to have accumulated deletions in the *mga* genes (Cleary *et al.*, 1998). Whilst the evidence points to HA production being regulated by other means than Mga (Falaleeva *et al.*, 2014; Velineni and Timoney, 2018), previously it has been observed that *S. pyogenes* strains maintained in prolonged stationary phase (3-5 days) present with ceased or reduced expression of multiple virulence factors

including Mga, M protein and capsule (Leonard, Woischnik and Podbielski, 1998). Mga is a transcriptional regulator and it has been observed that it possesses more than one DNA binding domain which has lead to speculation that it is capable on acting on numerous promoter targets (McIver and Myles, 2002). Whilst none of these observations indicate a direct link between capsule production and Mga regulation, the observation of this mutation between mucoid and non-mucoid phenotypes in conjunction with the accepted notion that capsule is a virulence factor could be considered cause for further investigation into the effects of Mga function on phenotype with respect to capsule. The majority of research into the *mga* operon and regulon has been conducted in *S. pyogenes* however a comparison of the operons indicates that the role of this regulator may differ slightly in *S. equi*.

3.4.3.3. RNA polymerase sigma-70 factor family

The missing stretch of sequence at the beginning of the gene in 40327NMV does not appear to affect either of the regions identified in pfam as promoter recognition or binding regions. Sigma factors provide yet another mechanism for pathogenic bacteria to respond to changes in environmental conditions such as pH and temperature and also have been found to have a role in the transcription of virulence factors (Opdyke, Scott and Morgan Jr, 2001). Members of the sigma-70 family of sigma factors associate with RNA polymerase in order to direct the enzyme to the relevant promoter sequences and thus is essential for transcription initiation (Paget and Helmann, 2003). Whilst they have been associated with regulation of virulence factors, sigma-70 family includes the “housekeeping” sigma factor as well as members specifically associated with stress responses (Kazmierczak, Wiedmann and Boor, 2005) and capsule production has been associated with response to stress stimuli such as pH changes and oxidative stress (Cleary and Larkin, 1979; Liu *et al.*, 2008; Henningham *et al.*, 2015; Chiang-Ni *et al.*, 2016). It has been claimed that a promoter motif for sigma-70 factor has been identified upstream of the *hasE/pgi* gene in *S. equi* species, speculating that this

gene is under the regulation of more than one operon due to its essentiality in other cellular processes (Blank, Hugenholtz and Nielsen, 2008). Thus it is not illogical to speculate that sigma factors play a role in regulation of capsule production. Whether the loss of sequence at the beginning of the RNA polymerase sigma-70 factor gene in 40327NMV plays a role in loss of HA production is uncertain since as previously mentioned the loss has not occurred within conserved regions. Further investigation in this case is outwith the scope and resources of this project.

3.4.4. The *hasC* paralogs

Due to the sequencing and annotation challenges described above both paralogs were not successfully identified from the RAST annotated genomes of the strains of interest apart from SEZ20727 and in some cases such as SEE20561 neither were identified. As such the *S. equi* subsp. *equi* reference strain 4047 (SE4047) was used as the complete genome of this strain is available. The *hasC* paralogs were previously observed and analysed in *S. equi* subspecies *zooepidemicus* and *equi* as part of a study analysing the evolution of the *has* operon (Blank, Hugenholtz and Nielsen, 2008). In this work it was observed that both *S. zooepidemicus* and *S. equi* HasC paralogs were identical apart from a portion of the sequences at the C terminal. While this study did confirm this for *S. equi* subsp. *zooepidemicus* strain MGCS10565 the same was not found for either ATCC 35246, SEZ20727 or SE4047.

3.4.4.1. *S. equi* subsp. *equi* 4047 HasC paralogs

As described above (3.3.4.) the paralogs of this strain were almost identical, however an amino acid substitution was observed at position 11 which introduces two new molecular groups to the protein. The substitution of amino acids within proteins can cause conformational changes which may destabilise the structure or prevent correct folding and modification of the molecule, or in the case of a substitution within an active site the accessibility and binding affinity of an enzyme may be hindered or altered (Teng *et al.*, 2010). The substitution observed between the paralogs of SE4047 are found in a predicted active site as annotated in NCBI. This could indicate a change in binding capacity or activity between the two enzymes. Whether and to what extent both paralogs are active in *S. equi* species has not been investigated and activity would need to be assessed in order to determine whether the amino acid substitution observed in this strain might have an effect on capsule production.

3.4.4.2. *S. equi* subsp. *zooepidemicus* strain HasC paralogs

The sequences of the paralogs in *S. equi* subsp. *zooepidemicus* strains were all identical when compared to each other at amino acid level. A comparison of the HasC1 paralogs revealed two amino acid substitutions between SEZ35246 and the other two strains. Whilst neither of these changes were located within active sites of the enzymes there is still a possibility that activity may be affected due to changes in protein stability and confirmation. Unlike for SE4047 in which the paralogs were almost indistinguishable at nucleotide level, a comparison of the nucleotide sequences of the *hasC1* and *hasC2* paralogs in *S. equi* subsp. *zooepidemicus* strains displayed a series of single base substitutions throughout the sequence length which were not found to be translated into amino acid changes. Previously it was suggested that continual homologous recombination occurred within the *S. equi* subsp. *zooepidemicus* paralogs (Blank, Hugenholtz and Nielsen, 2008), which may explain the variation in nucleotide sequences observed in this study. As with SE4047, the activities of either HasC paralog and whether they differ has not been investigated.

3.4.5. The *murA* paralogs

As will be discussed in more detail in Chapter 4, the enzyme UDP-N-acetylglucosamine enolpyruvyl transferase (MurA) is essential to the formation of peptidoglycan, a key component of the cell wall in bacteria (Zoeiby, Sanschagrín and Levesque, 2003). It is responsible for catalysing the first step in the pathway where it transfers an enolpyruvate residue from a phosphoenolpyruvate molecule to *N*-acetylglucosamine (NAG) (Eschenburg *et al.*, 2005). As a result it competes directly with the hyaluronan synthase enzyme for NAG and so is an enzyme of interest in this study. Two paralogs of MurA exist in gram positive bacteria (Kedar *et al.*, 2008). This was originally discovered in *S. pneumoniae*, where both paralogs were observed to be active to varying degrees as well as differentially susceptible to the antibiotic Phosphomycin, which is specific to this enzyme (Du *et al.*, 2000).

Both paralogs were identified in our strains of interest and when compared between species displayed a large degree of identity which is indicative of the enzyme essentiality. Comparing the *S. equi* paralogs to those of *S. pneumoniae* strain R6 displayed a high degree of similarity in amino acid sequences. The MurA1 paralogs possessed 80.19% identity to one another. Several amino acid substitutions were present and although none were in a reported active site, one substitution of a valine in SE40327 to an isoleucine in R6 occurred within a designated hinge site. Hinge sites within proteins are areas of flexibility and are often difficult to predict (Flores *et al.*, 2007). The MurA enzyme of *E. coli* has previously been described to undergo severe conformational changes when bound to its substrate as detected by Small-angle X-ray scattering (SAXS) and fluorescence spectroscopy (Zoeiby, Sanschagrín and Levesque, 2003). Whether a change of amino acid at a hinge site may affect the flexibility and ease of these conformational changes upon binding and thus effect the enzyme activity could be speculated. The MurA1 paralog in *S. pneumoniae* R6 was described as the more catalytically active of the two enzymes (Du *et al.*, 2000), however this may not necessarily be the case in *S. equi* and so cannot be inferred especially if the amino acid content of the hinge sites differ between these species. A comparison of the MurA2

paralogs between these two species shows homology of 79.76% and no amino acid substitutes within active or hinge sites.

The *murA* paralogs of SE40327 possess 47.47% homology and a comparison of the amino acid sequences indicates that amino acid differences exist within active sites. This observation may explain the differences in activity observed between the two enzymes. The sequences of the *murA* paralogs of *S. equi* were identified in order to facilitate the design of gene expression analyses which will be discussed in chapter 4. The observation that the enzymes differ in sequence from those of *S. pneumoniae* R6 indicates that assumptions on their expression and activity cannot be made from the data presented in that study.

3.4.6. Conclusions

The sequencing of the genome of SE40327 provided the means to compare between other strains of *S. equi* and confirm the identity of the strain. The genomes of SEE20561 and SEZ20727 were not previously available but now can be utilised in other genome comparison analyses. Whilst the separate sequencing of the mucoid and non-mucoid phenotypes of SE40327 did not result in the revelation of any obvious functionally relevant differences with respect to the genes directly involved in HA production or provide information on the *has* operon structure of the strains, it has provided some avenues for continued investigations. In addition the sequence information of the MurA enzymes in these strains may be useful for future analyses.

Suggestions for investigations and analyses going forward will be discussed in chapter 6.

4. Competing Metabolic Pathways of Hyaluronic Acid Synthesis

4.1. Introduction

There are two precursors involved in the synthesis of HA: UDP-Glucuronic acid and UDP-N-acetylglucosamine and both are derived from glucose through two separate pathways which involve conversion into glucose-6-phosphate and fructose-6-phosphate (see figure 40). Glucose-6-phosphate is also involved in the production of other cell wall components, lactic acid and accumulation of biomass. As previously discussed, it is converted to HA component UDP-Glucuronic acid (Sze, Brownlie and Love, 2016). Fructose-6-phosphate is converted through several steps into one of the precursors to HA, UDP-N-acetylglucosamine which itself is also utilised in the synthesis of peptidoglycan, the key component of the cell walls of gram positive bacteria (Yamada and Kawasaki, 2005). Thus HA production is constantly in competition with other necessary cellular processes. With regards to UDP-Glucuronic acid and UDP-N-acetylglucosamine, it has previously been reported that the latter is the pivotal precursor relating to increasing HA molecular weight, with increased concentrations of it associated with increased molecular weight (Chen *et al.*, 2009b). Following this it was discovered that this effect was only observed upon overexpression of *hasD* and *hasE*, the genes responsible for synthesising this precursor rather than with supplementation of the media with precursors themselves and that a cut off concentration exists where increasing expression past this does not result in further increases to molecular weight (Chen *et al.*, 2014). As a result of that research it was observed that the presence of the empty vector plasmid pNZ8148 also induced an increase in HA molecular weight and yield and upon investigation of this effect it was discovered that the cause could be attributed to upregulation of HasD which is responsible for the last steps of N-acetylglucosamine synthesis and unexpectedly, downregulation of the enzyme responsible for the first step of peptidoglycan synthesis MurA which is not directly involved in any stage of HA synthesis and directly competes with HasA for UDP-N-acetylglucosamine (Marcellin, Chen and Nielsen, 2010).

Peptidoglycan is a fundamental component of the bacterial cell wall which protects the cell from mechanical and osmotic stress and aids in maintaining cell shape (Zoeiby, Sanschagrin and Levesque, 2003; Hrast *et al.*, 2014). It is

essential to bacterial viability and so enzymes involved in its synthesis are often targets for development of new antibiotics (Eschenburg *et al.*, 2005). The first step of its synthesis involves the transfer of an enolpyruvyl component from phosphoenolpyruvate to UDP-N-acetylglucosamine which is catalysed by the enzyme UDP-N-acetylglucosamine 1-carboxyvinyltransferase (MurA) (Barreteau *et al.*, 2008). There are two copies of MurA in gram positive bacteria, thought to have arisen by gene duplication and both of which are active enzymes (Du *et al.*, 2000). The enzyme contains two globular domains with the active site lying between them. This site is open, but becomes closed upon binding of the substrates (Klein and Bachelier, 2006). The active sites include Cys115 which is thought to be involved in release of the product and can be substituted with an aspartic acid without loss of catalytic activity (Eschenburg *et al.*, 2005). Also important is Lys22 residue which is involved in the binding of phosphoenolpyruvate and the initiation of the conformational change of the enzyme (Samland, Amrhein and Macheroux, 1999). The Cys115 residue is the target for the only currently established antibiotic targeting MurA, Phosphomycin. This antibiotic works by binding covalently to the Cys115 residue and impeding the release of the reaction product thus halting the synthesis of peptidoglycan (Jukic, Gobec and Sova, 2019). Some species of bacteria such as *Mycobacterium tuberculosis* are naturally resistant to Phosphomycin which is due to the absence of a cysteine at residue 115 (Eschenburg *et al.*, 2005; Engel *et al.*, 2013). The other mechanisms of Phosphomycin resistance include mutations which impede the uptake systems of which there are two, the α -glycerolphosphate and glucose-6-phosphate uptake systems. Mutations in these systems prevent entry of the antibiotic into the cell. Additionally, upregulation of *murA* expression promotes resistance (Michalopoulos, Livaditis and Gougoutas, 2011; Silver, 2017). There are investigations ongoing to identify alternative MurA inhibitors to Phosphomycin such as terreic acid or the PEP 1354 peptide, however for now it remains the only confirmed antibiotic with specificity to MurA (Molina-López, Sanschagrin and Levesque, 2006; Abbott *et al.*, 2010; Han *et al.*, 2010).

As the reduction of MurA was associated with increased HA molecular weight this work tested the hypothesis that inhibition of MurA activity by Phosphomycin at sub-lethal levels might result in increased availability of

UDP-N-Acetylglucosamine for HA synthesis by HasA and subsequently increase the yield and/or molecular weight of the HA produced. There are no EUCAST (<http://www.eucast.org>) minimum inhibitory concentration (MIC) breakpoint data available for Phosphomycin for *Streptococcus spp.* of any strain as it is not recommended for use in this species of bacteria. Marcellin, Chen and Nielsen (2010) reported that concentrations of Phosphomycin as low as 0.5 ng/mL were lethal to *S. zooepidemicus* ATCC 35246 and so did not report on the effects of antibiotic supplementation on HA production and molecular weight. The set of experiments described in this chapter aimed to assess the MICs of Phosphomycin for the *S. equi* mucoid and non-mucoid strains of interest in this study and if possible assess any effects on HA production induced by the presence of Phosphomycin in the culture media.

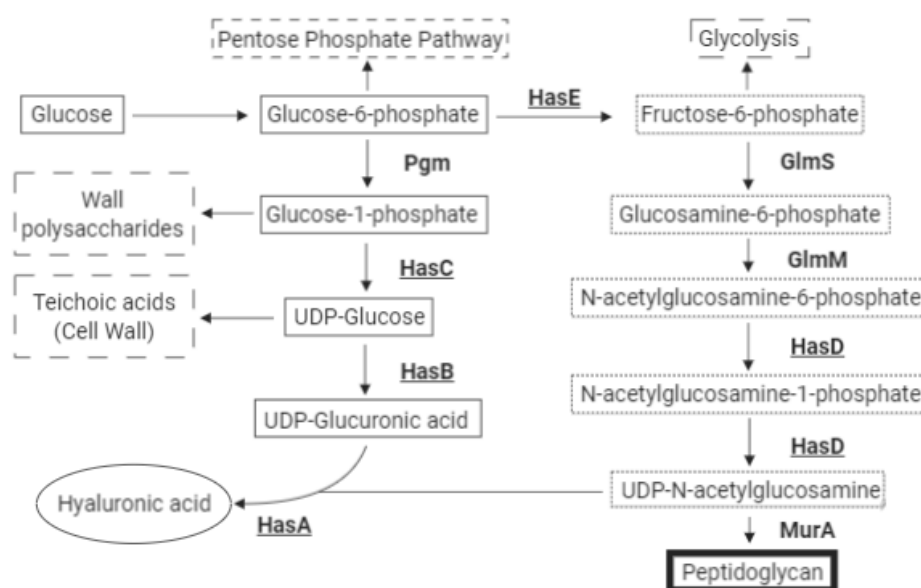


Figure 40 Hyaluronic acid production pathway precursors in order of synthesis with enzyme designations. The figure also displays pathways which are competing for precursors with Hyaluronic acid synthesis. Enzymes found within the *has* operon are underlined.

4.2. Materials and Methods

4.2.1. Preparation of antibiotic stocks

4.2.1.1. Phosphomycin

Phosphomycin disodium salt was obtained from Sigma Aldrich. Stock solutions of 1mg/mL (5.5 mM) were prepared in deionised water and sterilised. The stock solution was then stored refrigerated for up to 2 weeks.

4.2.1.2. Ampicillin

Ampicillin sodium salt was obtained from Sigma-Aldrich. Stock solutions of 1mg/mL (2.7 mM) were prepared by dissolving the solid in deionised water and sterilising by syringe filtering through 0.22 µm into aliquots. The aliquots were stored at -20 °C and defrosted at room temperature prior to use in minimum inhibitory concentration assays.

4.2.2. Antibiotic Minimum Inhibitory Concentration determination

Growth curves in the presence of antibiotic were analysed in accordance with EUCAST guidelines for determining minimum inhibitory concentrations of antibiotics (<http://www.eucast.org>), where the MIC of an antimicrobial substance is identified as the lowest concentration that inhibits visible bacterial growth. Bacterial growth in broth microdilution assays is observed as either turbidity or an accumulation of cells at the bottom of the wells depending on the type of organism.

Bacterial stocks were inoculated into 50 mL of BHI media which was pre-equilibrated to 37°C. The inoculum was then incubated overnight at 37°C with shaking at 150 RPM. A 96-well plate was seeded in columns with 100 µL of Brain Heart Infusion media. The layout of the plate is observed in figure 41. The maximum concentration of antibiotic to be analysed was prepared by diluting the stock in BHI media to the desired concentration and pre-equilibrating to 37°C. Antibiotic in BHI solution was seeded 200 µL into the empty column. Serial dilutions were performed by drawing 100 µL and mixing into the adjacent well, one column was left antibiotic free. Inoculum optical density was measured at 660 nm and corrected to 0.05 before 100 µL was seeded into each well with the blank column left

uninoculated. The plates were sealed with Breathe-Easy® sealing membranes and placed in the Omega FLUOstar plate reader (BMG Labtech) and incubated for 20 hours at 37°C. Optical density measurements at 660 nm were made every 30 minutes following 200 RPM shaking for 1s. Concentrations of Phosphomycin at which final OD was significantly different from the control was determined by One-Way ANOVA analyses and concentrations which differed significantly between each other were determined by Tukey test.

	1	2	3	4	5	6	7	8	9	10	11	12
A	B	N	0.0293	0.0586	0.1172	0.2344	0.4688	0.9375	1.875	3.75	7.5	15
B	B	N	0.0293	0.0586	0.1172	0.2344	0.4688	0.9375	1.875	3.75	7.5	15
C	B	N	0.0293	0.0586	0.1172	0.2344	0.4688	0.9375	1.875	3.75	7.5	15
D	B	N	0.0293	0.0586	0.1172	0.2344	0.4688	0.9375	1.875	3.75	7.5	15
E	B	N	0.0293	0.0586	0.1172	0.2344	0.4688	0.9375	1.875	3.75	7.5	15
F	B	N	0.0293	0.0586	0.1172	0.2344	0.4688	0.9375	1.875	3.75	7.5	15
G	B	N	0.0293	0.0586	0.1172	0.2344	0.4688	0.9375	1.875	3.75	7.5	15
H	B	N	0.0293	0.0586	0.1172	0.2344	0.4688	0.9375	1.875	3.75	7.5	15

Figure 41 Layout of the 96-well plate for Phosphomycin MIC determination. The antibiotic concentrations in each column are in µg/mL. A blank column (B) of uninoculated media and a negative (N) column of bacteria in media without antibiotic were also seeded to provide controls.

4.2.3. Capsule production analysis

Inoculum was prepared as previously described (2.2.2.1.) and used to prepare and culture 96-well plates as previously detailed (2.2.2.2). Bacteria were grown in BHI media supplemented with the chosen concentration of phosphomycin in order to assess any effects on growth and HA production. Following growth, samples for HPLC analyses were prepared as previously detailed (2.2.5.) and stored frozen at -20 °C until analysis.

4.2.4. RNA extraction

Bacteria stocks of 40327MV and 40327NMV were streaked onto SLA plates and incubated at 37°C for 48 hours. Single colonies were selected and inoculated into 10 mL of BHI media and grown at 37°C with 150 RPM shaking overnight before 1 mL of inoculum was added to 10 mL of BHI and grown at 37°C and 150 RPM shaking for 8 hours. Following this time 40327MV was treated with 10 mL 3.6 mM hyaluronidase (Sigma-Aldrich) in BHI media and incubated for 20 minutes at 37°C. The cultures were centrifuged at 4969 x g for 10 minutes, the pellets were washed 3 times in PBS (Gibco™) and stored at -80°C until they were freeze dried. Five 2.4

mm metal beads (Fisherbrand™) were added to each tube of freeze dried pellets and the tubes placed in a Precellys Evolution Homogenizer (Bertin Instruments) and beaten at 4500 RPM for 30 seconds. To extract RNA 500 µL of Trizol reagent (Ambion®) was added to each tube on ice and mixed by inversion before being incubated for 5 minutes at room temperature. Following incubation 100 µL of chloroform was added to each tube and the tubes shaken vigorously for approximately 15 seconds then incubated at room temperature for 5 minutes. Samples were centrifuged at 12 000 x g for 15 minutes at 4°C to separate into phases. The upper aqueous phase was removed by pipette and transferred to a fresh RNase free Eppendorf tube without contaminating with the intermediate phase, 500 µL of isopropanol was added to each sample of aqueous phase and the tubes inverted several times to mix before being incubated at room temperature for 10 minutes. The tubes were centrifuged at 12 000 x g for 10 minutes at 4°C. The supernatant was removed and 500 µL 75% ethanol was added. Samples were briefly vortexed and centrifuged at 7500 x g for 5 minutes at 4°C. The ethanol was removed and the tubes left horizontally to air dry. Once visibly dry 50 µL of nanopure water (Hyclone™) was added to each tube and mixed thoroughly by pipette before being placed in incubation at 55°C for 10 minutes to aid dissolution, the samples were then placed on ice prior to DNase treatment. All reagents for DNase treatment were sourced from ThermoFisher Scientific. To each tube, 5 µL of 10X reaction buffer with MgCl₂ was added followed by 2 µL of RNase free DNase. Samples were incubated at 37°C for one hour. To inactivate the DNase 5 µL of 50 mM EDTA was added and the tubes were heated to 65 °C for 10 minutes. To re-extract the RNA 5 µL of 3M sodium acetate solution (Sigma-Aldrich) was added, followed by 65 µL of isopropanol. Samples were incubated at -20°C overnight before being centrifuged at 20 000 x g for 30 minutes at 4°C. The supernatant was removed without disturbing the pellet and 500 µL ice cold 70% ethanol was used to wash the pellet, the tubes were vortexed briefly and centrifuged at 20 000 x g for 10 minutes at 4°C. The ethanol was removed and the tubes left horizontally to air dry. Once visibly dry 50 µL of nanopure water was added and the tubes incubated at 55°C for 10 minutes. The tubes were stored on ice until purity measurements were made with the NanoDrop™ 2000. Acceptable purity

criteria were a minimum of 100 ng/ μ L, a 260/280 ratio of 1.8-2.2 and a 260/230 ratio of 1.5-2.4. The RNA samples were then stored at -80°C until required for use where they were defrosted and stored on ice. Samples were run on a 2% agarose gel with SYBR[®] Safe DNA gel stain (Invitrogen) in order to confirm purity and samples were included or excluded from cDNA synthesis and subsequent PCR based on this result.

4.2.5. cDNA synthesis

The synthesis of cDNA from the RNA samples was carried out using the High-Capacity RNA-to-cDNA[™] kit from Applied Biosystems. Reactions were set up and carried out as per the manufacturer's instructions on ice. The reaction mix was aliquoted into Eppendorf tubes and centrifuged briefly to remove any bubbles and to mix contents. The tubes were then placed in incubation at 37°C for one hour. The tubes were then wrapped in parafilm and placed in a water bath at 95°C for 5 minutes to inactivate the reverse transcriptase enzyme before being frozen at -20°C until required for use in PCR reaction.

4.2.6. Primer design for RT PCR

Primers were designed using Primer3Plus (<http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>) with the following criteria. Primers must be 20-24 bases in length, product size must be 80-120 bases in length and the GC% must be 40-50. The optimum annealing temperature (T_m) was 60 °C. Primers were designed with G or C as the final nucleotide at the 3' end. Two housekeeping genes were used, the first was NAD-dependent glyceraldehyde-3-phosphate dehydrogenase (*plr*) and the second DNA gyrase subunit A (*gyrA*). The genes of interest investigated were *hasA*, *murA1* and *murA2*. The final sets of primers designed for the genes of interest are presented in Table 10. Primers were ordered from Eurofins (<https://www.eurofinsgenomics.eu>) as custom DNA Oligos in tubes and were reconstituted and stored as per manufacturer's instructions.

gene name	Forward 5'-3'	Reverse 5'-3'
<i>gyrA</i>	TGACAAGGACAAGGCAGAAG	CATTATCCGCAATAGCGAGAC
<i>hasA</i>	AGGCCCACTTAGCGTTTACA	GATGACAGGTGCTTGACCAAC
<i>murA1</i>	TGGGGTAAGGATTACCAATGTC	AATGGGTGTGAGAATGACGG
<i>murA2</i>	TGACAAGGACAAGGCAGAAG	CATTATCCGCAATAGCGAGAC

Table 10 Primers designed using Primer3Plus for the respective genes of interest. The primers were designed and used for real time PCR to determine transcript levels of the respective genes of interest in mucoid and non-mucoid phenotypes.

4.3. Results

4.3.1. Phosphomycin MICs of *S. equi* strains of interest

In order to determine the Phosphomycin MIC for each *S. equi* strains the bacteria were cultured in varying concentrations of the antibiotic in BHI media. Conventionally, MICs are assessed visually at a growth endpoint (16-20 hours) – in this instance, growth was monitored continually and the results are presented in the following graphs. Figure 42 displays the growth results for SE40327. A dramatic reduction in growth is observed at a Phosphomycin concentration of 15 µg/mL, however the reduction in final OD reached seen at concentrations of 7.5, 3.75, 1.875 and 0.9375 µg/mL is significant also ($p = 0.000$). Final OD at concentrations 0.9375 and 1.875 µg/mL were not significantly different from each other whereas concentrations 3.75, 7.5 and 15 µg/mL were all significantly different ($p = 0.000$). Lower concentrations than this did not result in a significant reduction in final OD or a visible absence of growth compared to the control. Turbidity was observed for all concentrations apart from 15 µg/mL following the incubation period.

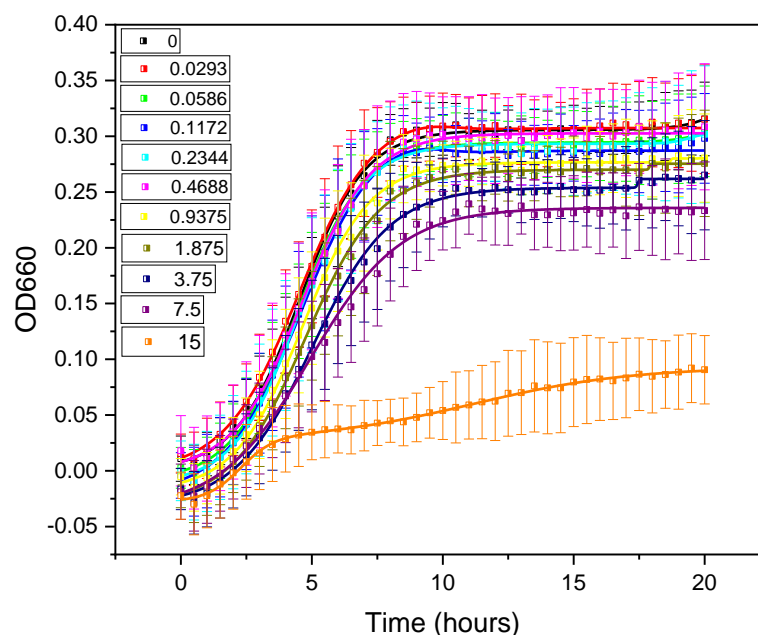


Figure 42 Phosphomycin MIC of SE40327 in BHI media. Concentrations of Phosphomycin are in $\mu\text{g/mL}$. The curves plotted are optical density at 660 nm over time. Points on the graph are representative of mean values over 3 biological replicates and error bars are standard deviation.

Figure 43 displays the MIC results for SEE20561. Most noticeable is the absence of growth observed at 15 $\mu\text{g/mL}$, although there is an apparent increase in OD in the first 4 hours according to the graph this did not result in visible turbidity following 20 hours of incubation. A significant reduction in final OD was observed at 7.5, 3.75 and 1.875 $\mu\text{g/mL}$ compared to the control ($p = 0.000$) and the final OD for each of these concentrations also differed significantly from each other. Visible turbidity was observed following 20 hours of incubation for all concentrations bar 15 $\mu\text{g/mL}$.

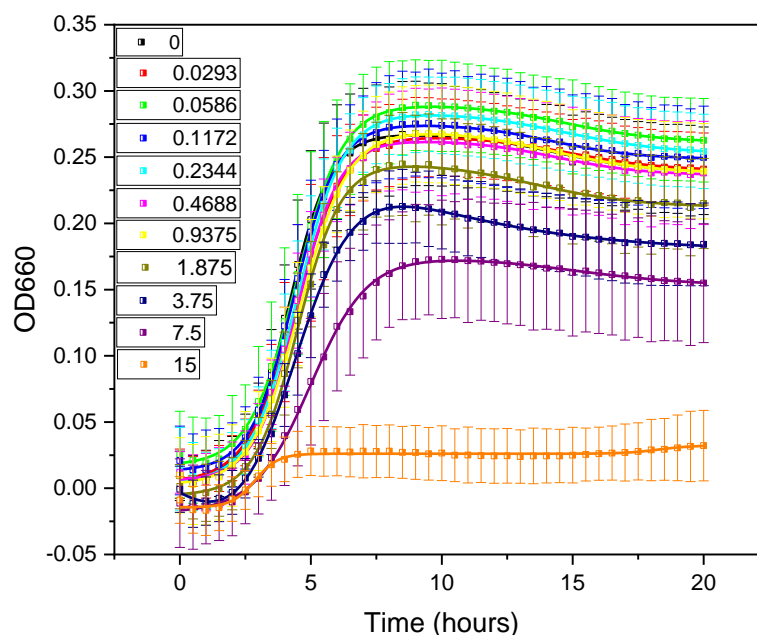


Figure 43 Phosphomycin MIC of SEE20561 in BHI media. Concentrations of Phosphomycin are in µg/mL. The curves plotted are optical density at 660 nm over time. Points on the graph are representative of mean values over 3 biological replicates and error bars are standard deviation.

Figure 44 shows the MIC results for SEZ20727. Once again there is an evident inhibition of growth at 15 µg/mL with any apparent increase in OD observed on the graph being indiscernible by eye following 20 hours of incubation. Growth of this strain in 7.5 µg/mL Phosphomycin appears to be variable with a reduced average maximum optical density but large error bars indicating variance between biological replicates at this concentration. Significant differences in growth compared to the control were observed at concentrations as low as 0.0293 µg/mL ($p = 0.000$) however wells were observed to be turbid following the incubation period at all concentrations apart from 15 µg/mL demonstrating that these concentrations were not inhibitory according to EUCAST guidelines.

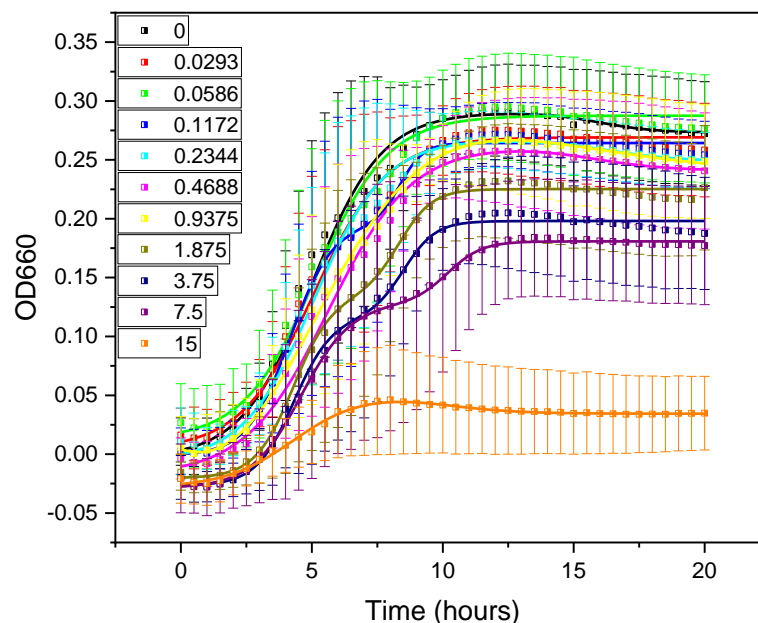


Figure 44 Phosphomycin MIC of SEZ20727 in BHI media. Concentrations of Phosphomycin are in $\mu\text{g/mL}$. The curves plotted are optical density at 660 nm over time. Points on the graph are representative of mean values over 3 biological replicates and error bars are standard deviation.

Figure 45 displays the MIC results for SER17037. Inhibition of growth was evident by lack of visible turbidity at 15 $\mu\text{g/mL}$ as observed with the other strains and this appears to be supported by the curve in the graph. Another pronounced feature of the graph is that while inhibition of growth is observed at 7.5 $\mu\text{g/mL}$ the size of the error bars indicate considerable variation between biological replicates with regards to this concentration. When manually reading the MIC's in some of the biological replicates turbidity was prominent in the 7.5 $\mu\text{g/mL}$ whereas in others there was barely visible turbidity, this variation is supported by the large error bars. The final OD's were significantly different to the control for Phosphomycin concentrations as low as 0.0293 $\mu\text{g/mL}$ however growth was visually observed for all concentrations below 15 $\mu\text{g/mL}$. Unexpectedly, no significant difference was found between final OD of the control and that of concentration 0.2344 $\mu\text{g/mL}$ despite the values at concentrations below this being significantly decreased compared to the control ($p = 0.000$). An additional observation was that the shape of the growth curves at concentrations 1.875-7.5 $\mu\text{g/mL}$ appeared different from

the others with what can be described as a ‘dip’ occurring in the latter half of log phase.

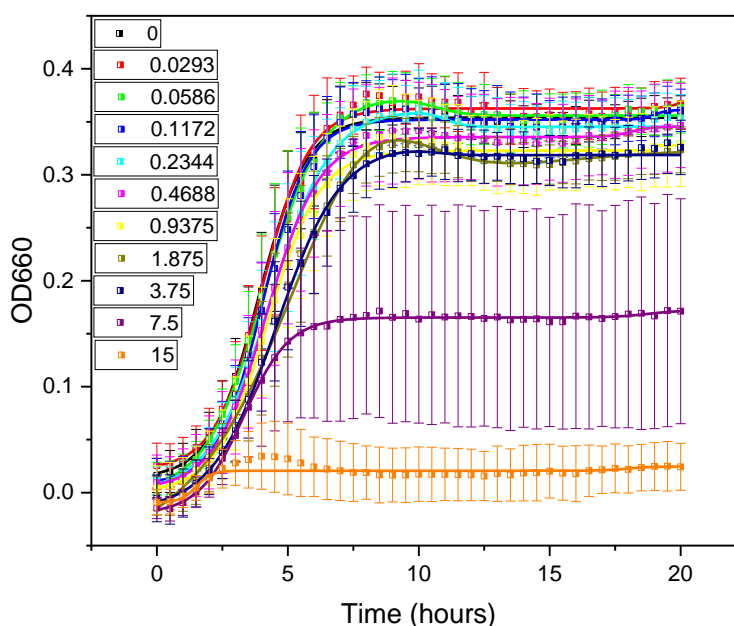


Figure 45 Phosphomycin MIC of SER17037 in BHI media. Concentrations of Phosphomycin are in $\mu\text{g/mL}$. The curves plotted are optical density at 660 nm over time. Points on the graph are representative of mean values over 3 biological replicates and error bars are standard deviation.

From the results of the growth experiments in varying Phosphomycin concentrations it can be deduced that the minimum inhibitory concentration in BHI media is $15 \mu\text{g/mL}$. This is in accordance with the MIC criteria from EUCAST which dictates that the lowest concentration which results in no visible growth of an organism may be recorded as the MIC. The lack of visible growth in the wells containing $15 \mu\text{g/mL}$ Phosphomycin were confirmed upon analysis of growth curve data which showed considerably reduced growth curves. In addition to this result, the observation that the final OD reached is significantly different at concentrations much lower than the inhibitory concentration when compared to the control provides useful information as to the effect of Phosphomycin supplementation at below growth inhibitory concentrations. The lack of visible growth inhibition at concentrations of $7.5 \mu\text{g/mL}$ indicates that growth in the presence of $\mu\text{g/mL}$ concentrations of Phosphomycin is possible for all 4 strains of interest.

4.3.2. Capsule production in the presence of Phosphomycin

In order to investigate the effects of Phosphomycin on capsule production, the bacteria were cultured in the presence of 2 and 7.5 $\mu\text{g/mL}$ ($1/2$ and $1/12^{\text{th}}$ MIC) of the antibiotic and the HA concentration of the culture media measured by HPLC. As described previously, the concentrations from each biological replicate were corrected to the maximum OD achieved in order to normalise the data. The HA retention time was slightly higher than in BHI media without Phosphomycin with the average retention time being 4.1 minutes compared to 3.9 minutes in BHI alone. The results of the HPLC analyses are displayed in table's 11 and 12. HA concentration whilst appearing relatively consistent also appears to be reduced when compared to the concentration in BHI media alone (Tables 7 and 8, Chapter 2).

The data was analysed by One-Way ANOVA, the results of which were as follows, SE40327 $df = 1, 4, f = 0.90, p = 0.396$ and SER17037 $df = 1, 4, f = 0.07, p = 0.798$. For both strains there was no significant difference in HA production between Phosphomycin concentrations 2 and 7.5 $\mu\text{g/mL}$ as indicated by the p values being greater than 0.05. The results were then compared to the HA production results in BHI media alone (Tables 7 and 8 Chapter 2.3.4) for each strain. A significant reduction in HA concentration was observed when SE40327 was cultured in BHI containing either concentration of Phosphomycin compared to that of BHI media alone (One-Way ANOVA, $df = 2, 6, f = 21.14, p = 0.002$), the same was observed for SER17037 (One-Way ANOVA, $df = 2, 6, f = 11.09, p = 0.010$).

SE40327				
Phos conc. $\mu\text{g/mL}$	Retention time (min)	HA concentration ($\mu\text{g/mL}$)	Final OD660	Corrected concentration to max OD ($\mu\text{g/mL}$)
2	4.095	52.05	0.46	56.07
2	4.194	60.62	0.5	60.62
2	4.105	49.61	0.25	100.84
7.5	4.171	69.34	0.52	69.34
7.5	4.116	48.28	0.06	602.11
7.5	3.86	52.55	0.47	58.62

Table 11 Hyaluronic acid concentration produced by SE40327 in BHI media supplemented with Phosphomycin. The table displays the results of HPLC analyses on cell stripped media following 8 hours of growth in BHI supplemented with either 2 or 7.5 $\mu\text{g/mL}$ Phosphomycin. Final optical density values are displayed along with the measured hyaluronic acid concentration and retention time for each analysed sample. Concentrations were corrected to the maximum OD

reached for each media. Each entry represents a single biological replicate of 15 pooled technical replicates.

SER17037				
Phos conc. µg/mL	Retention time (min)	HA concentration (ug/ml)	Final OD660	Corrected concentration to max OD (µg/ml)
2	3.985	66.72	0.45	77.13
2	4.229	61.83	0.47	67.88
2	4.215	66.82	0.52	66.82
7.5	4.113	62.63	0.4	71.09
7.5	4.174	70	0.45	70
7.5	4.158	60.92	0.41	67.95

Table 12 Hyaluronic acid concentration produced by SER17037 in BHI media supplemented with Phosphomycin. The table displays the results of HPLC analyses on cell stripped media following 8 hours of growth in BHI supplemented with either 2 or 7.5 µg/mL Phosphomycin. Final optical density values are displayed along with the measured hyaluronic acid concentration and retention time for each analysed sample. Concentrations were corrected to the maximum OD reached for each media. Each entry represents a single biological replicate of 15 pooled technical replicates.

4.3.3. RT-PCR analysis of *murA* paralog gene expression between phenotypes

Attempts to develop RT-PCR to determine any differences in expression of *hasA*, *murA1* and *murA2* between mucoid phenotypes were made. RNA isolation and purification was carried out as described (4.2.4.) Despite multiple attempts to optimise all stages such as RNA purification, primer design, template and RT enzyme concentration and RT-PCR method, it was not possible to obtain sufficient or consistent results. Specific aspects in this development and mitigation approaches are discussed below (4.2.3 and following sections).

4.4. Discussion and Conclusions

As previously mentioned, UDP-N-acetylglucosamine is considered the pivotal precursor for molecular weight control of HA produced by *S. equi* subsp. *zooepidemicus*. This precursor is also used for the synthesis of peptidoglycan which is a crucial component of bacterial cell walls (Barreteau *et al.*, 2008). UDP-N-acetylglucosamine is the substrate for the enzyme involved in the first step of this pathway MurA and as such this enzyme competes directly with HasA. Due to this direct competition it was theorised that reducing the action of MurA would increase the availability of this precursor for HA synthesis.

The study by Chen *et al.* (2010) identified that a reduction in MurA enzyme was associated with an increase in HA molecular weight produced by *S. equi* subsp. *zooepidemicus* ATCC 35246 and this was thought to be due to the fact MurA competes directly with HasA for the HA precursor UDP-N-acetylglucosamine. Phosphomycin is an antibiotic in which the mechanism of action is inhibition of MurA specifically. It was hypothesised that supplementation of culture media with Phosphomycin at a less than inhibitory concentration may reduce the activity of MurA and as a result increase the availability of the precursor for HA production. The first step to assess this was to determine the MIC's of Phosphomycin in the strains of interest. Following this HA production in sub-inhibitory concentrations was assessed.

4.4.1. Phosphomycin MIC's in *S. equi* strains

Marcellin *et al.* (2010) reported that Phosphomycin concentrations as low as 0.5 ng/mL were lethal to the *S. zooepidemicus* strain of interest in their study. This indicates extremely potent antimicrobial activity which is more typically in the µg/mL range. Phosphomycin MICs for other gram positive species such as methicillin susceptible *S. aureus* strains were found to range 0.5-32 µg/mL with methicillin resistant strains being as high as 128 µg/mL and in the same study overall from the 590 common bacterial species tested the Phosphomycin MICs varied 0.25-512 µg/mL (Lu *et al.*, 2011). Due to the variation in MIC's observed between strains the effects of Phosphomycin on growth of the strains of interest in this study were investigated as it was postulated that they may not be as low as claimed by Marcellin *et al.* (2010). Phosphomycin is transported into the cell via two uptake systems, the first is the L-α-glycerophosphate uptake system and the second is the glucose-6-phosphate (G6P) uptake system (Zoeiby, Sanschagrín and Levesque, 2003). Due to this, when testing the susceptibility of Phosphomycin the media is often supplemented with G6P in order to enhance the antimicrobial activity (Lu *et al.*, 2011; Michalopoulos, Livaditis and Gougoutas, 2011; Falagas *et al.*, 2016). However for the purpose of this study it was not necessary to do this as the ultimate aim was to investigate the possibility of introducing Phosphomycin into the current fermentation process for HA production in our strains of interest. From the results it can be concluded that the

visually observed MIC for all four strains is 15 µg/mL in BHI media, a large contrast to the reported lethal concentration of 0.5 ng/mL observed for *S. zooepidemicus* ATCC 35246. No EUCAST breakpoint data for *Streptococcus* groups A, B, C and G are available as these organisms are not recommended targets for Phosphomycin treatment, however it has been claimed MIC values for Phosphomycin of ≤ 64 µg/mL are considered susceptible (Michalopoulos, Livaditis and Gougoutas, 2011). Accepting this criteria it can be concluded that all four strains of interest are susceptible to Phosphomycin in BHI media. Growth curves are not a EUCAST requirement for susceptibility testing however they were performed in order to provide a range of concentrations which did not inhibit growth for the purpose of this project. Due to the previously reported low lethal concentration of Phosphomycin on *S. zooepidemicus* MIC assays were originally set up with a maximum concentration of 0.5 µg/mL Phosphomycin however no effect on growth was observed so the maximum concentration was increased to 1, 3, 5 and 10 µg/mL before arriving at the concentration of 15 µg/mL (Data not shown). In addition, MIC's of Ampicillin were determined for each strain as a control. Breakpoint data from EUCAST declare that MIC's for Ampicillin in group A, B, C and G *Streptococcus* of ≤ 0.25 µg/mL indicate susceptibility. Inhibition of growth was observed manually at 0.0156 µg/mL for SE40327 and SER17037, 0.0625 µg/mL for SEE20561 and 0.0313 µg/mL for SEZ20727 indicating that all strains were susceptible to Ampicillin in BHI (Data not shown). As previously mentioned the aim was to establish a range of below inhibitory concentrations of antibiotic to investigate further.

4.4.2. Capsule production in the presence of Phosphomycin

It has previously been reported that N-acetylglucosamine (NAG) is the limiting precursor with regards to HA molecular weight (Chen et al., 2009; Chen et al., 2014). As Phosphomycin specifically targets MurA which directly competes with HasA for NAG, it was hypothesised that supplementing the media with a non-inhibitory concentration of the antibiotic would result in an increase in HA production and or molecular weight as it would increase NAG availability. The highest concentration of Phosphomycin which did not result in visible inhibition of growth in the

strains was 7.5 µg/mL and so this was chosen as a concentration for analysis of HA production with Phosphomycin. With regards to SER17037 the large error bars observed for 7.5 µg/mL were considered an indication of variation between biological replicates, possibly due to the MIC being nearer to 7.5 than 15 µg/mL, however as turbidity was manually observed in all replicates this concentration was not considered inhibitory by EUCAST guidelines and so was tested in SER17037 regardless of the variation observed in the growth curve. Since an aim was to assess the highest Phosphomycin concentration possible without inflicting a marked change to the growth of the strains, a concentration of 2 µg/mL was selected in addition. This was chosen after analysing the growth curve data the concentration of 1.875 µg/mL was considered to be the highest concentration reached before the curves differed markedly compared to the control curve. Bacteria cultures were set up and grown as previously described (2.2.5) in order to ensure results would be comparable to non-antibiotic supplemented growth.

From the results of the HPLC analyses this was not the case however with both capsule producing strains displaying significantly reduced HA concentrations compared to when cultured in BHI media alone. In addition to the reduced concentration, although the HPLC was not set up to measure molecular weight the slightly increased retention time may indicate that the HA molecular weight is also reduced compared to that produced in BHI media alone.

The reduced HA production in the presence of Phosphomycin was an unexpected result. It is possible that this observation is the result of the emergence of a tolerance strategy in these strains. Mechanisms of resistance to Phosphomycin include mutational changes in either MurA which lowers the affinity to the antibiotic or in the uptake systems which reduces entry of Phosphomycin into the bacterial cells. Additionally, increasing MurA expression is another means of developing resistance (Falagas *et al.*, 2016; Silver, 2017). It was previously reported that over expression of either of the copies of *murA* lead to a reduction in HA molecular weight (Marcellin, Chen and Nielsen, 2010) and so it is possible

that to counteract the presence of Phosphomycin in the media, expression of *murA* is increased resulting in reduced HA production.

4.4.3. Gene expression of the *MurA* paralogs between phenotypes

In gram positive bacteria there are two paralogs of the MurA enzyme which compete with HasA for UDP-N-acetylglucosamine. Both paralogs were found to be active to different degrees and both are susceptible to inhibition by Phosphomycin (Du *et al.*, 2000). While it has been described that the paralogs have different levels of activity, there has been no consideration as to whether the genes are expressed concurrently and – if so - at the same level and whether changes in expression occur between mucoid and non-mucoid phenotypes. In order to assess this attempts were made to measure expression levels of the genes during log phase using RT-PCR.

RT-PCR is a commonly used method for analysing gene expression as it utilises mRNA as the template thus giving a representation of the levels of a target gene within a sample (Wacker and Godard, 2005). It can be carried out in either one step or two step reactions and involves the conversion of the mRNA template to cDNA prior to amplification in the PCR reaction (Jalali, Zaborowska and Jalali, 2017)

4.4.3.1. Primer design

As previously mentioned primer design was carried out using Primer3Plus software. The criteria for primers was product size 80-120 bases, primer length 20-22 bases, GC content 40-60%, and containing a G or C at the 3' end of the sequence. In addition primers with three or more repeats of a single nucleotide within were avoided where possible to reduce instances of primer dimers or hair pins. Melting temperatures (T_m) of 57-62° were considered acceptable. All primers were checked for specificity using NCBI BLAST and by searching the nucleotide sequences in RAST genome annotations.

The housekeeping gene selected was DNA gyrase A (*gyrA*) and primer sequences for this gene detailed previously in *S. equi* subsp *equi* were used (Steward, Robinson and Waller, 2016). *Streptococci*

possess low GC content in their genomes (33.75-43.40%) (Gao *et al.*, 2014). This causes some difficulties in primer design as ideally the GC content of primers and amplicons should be as close to 50% as possible to ensure stability (Bustin and Huggett, 2017). The genes of interest (*murA1*, *murA2* and *hasA*) were 39-47% GC which meant designing primers with a desirable content and containing minimal base repeats was difficult as often the more GC rich areas of the sequences contained a large number of repeats. Whilst Primer3Plus was a useful tool for selecting areas of the genes where sequences might be suitable for primers, the majority of the initial primers suggested were not considered suitable, either due to similarities in sequences between the forward and reverse primers which was predicted to increase the chances of primer dimer formation, or due to containing an unacceptable number of repeats which was predicted to increase the chances of hair pins forming. In the case of *hasA* three sets of primers were designed and obtained and following testing the second set was selected as the final primers for PCR. The *murA* paralogs provided additional complication. The nucleotide sequences are 53.9% identical as calculated by EMBOSS needle (<https://www.ebi.ac.uk/Tools/psa/>) and have very similar GC content. Areas of the sequences which possessed less identity were AT rich making them less suitable for primer design. The similarity of the two sequences to each other meant that designing primers which would be specific to the single paralogs was complicated as there was an increased risk of nonspecific annealing and double PCR product formation. As a result two sets of primers were designed for each paralog. The first primer set designed and ordered for *murA1* were specific for the target as indicated by single peaks in the melt curve. However the first primer set for *murA2* appeared to not be specific enough under the original conditions tested and so a second set was designed. Another issue identified with the use of Primer3Plus for primer design was that the predicted T_m of the primers was often not the same as that detailed on the product report provided by the manufacturer and so the T_m of the primers for each gene of interest differed to a greater extent than was predicted during their design.

However all fell within the accepted range and so attempts to design the assay were continued.

In addition to the previously mentioned genes of interest primers for additional genes were designed to investigate their expression within the two phenotypes, however in practice these primers were not suitable as they were not specific enough and did not result in amplification of the target. As a result these were omitted from the development of the PCR assay with the aim of returning to the targets in future. However, investigation of expression was continued with the housekeeping gene *gyrA* and the three main genes of interest *hasA*, *murA1* and *murA2* as these were the most important to the questions being posed.

In conclusion the primer design for the genes of interest in *S. equi* subsp *equi* 40327 was not straight forward using the criteria selected due to both the low GC content of the genes as well as a large number of areas of repeats. In addition the design of specific primers for *murA1* and *murA2* is challenging due to great similarity between the nucleotide sequences. For *murA2* the ideal primers fell within an area of sequence which is too similar to *murA1* to be confident that expression of this paralog alone is being measured.

4.4.3.2. PCR assay optimization

Initially, gene expression between phenotypes was attempted to be measured using the RNA extracted and purified and the Power SYBR® Green RNA-to-CT™ 1-Step Kit (Thermofisher) as this was available within the lab and had been successfully utilised for gene expression analyses in *E. coli*. One-Step RT-PCR involves performing the reverse transcription (RT) and subsequent PCR reaction within the same tube, resulting in reduced handling of samples and a simplified set up which may reduce chances of contamination (Wacker and Godard, 2005). However for the purpose of this study the results from One-Step PCR method overall lacked the desired precision with replicates being inconsistent in the majority of targets. A drawback of the One-Step RT-PCR method is that

optimized mixtures are provided in kits which can be costly and are often optimized using only one target gene and one template source, in the case of the mastermix used in this study it appears that optimisation was carried out on the gene for GAPDH in the human kidney cell line BOSC23 which has a GC content of 59%. In addition, both reactions occurring in the same tube reduces the likelihood of the conditions being optimal for either single reaction and troubleshooting is hindered due to a lack of control of the sequential steps (Prediger, 2012). One possible reason for the variation was that all target genes are present in the genome in a single copy and so the initial quantity of 5 ng RNA did not provide enough starting material. The kit used was compatible with a maximum concentration of 100 ng RNA and so the concentration used was increased to 10 and then 50 ng both with and without adjusting the RT enzyme concentration. However this did not decrease Ct values as anticipated, indicating that some form of inhibition may be present. The RNA quantity and purity had previously been measured following extraction using the Nanodrop™ 2000 and so further analyses of purity were carried out where the samples were all run on an agarose gel (Results in appendix B). None of the samples appeared to be significantly degraded, however in some of the samples smearing was observed and the presence of some form of contaminant remained in the wells, this was speculated to be ethanol contamination from the final RNA pellet wash. Following this only the samples which did not have evidence of any contamination were used. Another assay was performed in which a series of dilutions (1/2, 1/5 and 1/10) of 10 ng of the RNA was used in the reactions. The results demonstrated similar amplification curves for all template dilutions which suggested inhibition was present within the reaction. One-step RT-PCR is considered to be less sensitive than two-step due to the inability to optimise both reactions independently of each other (Bioline, 2015). In particular it can be difficult to troubleshoot the RT reaction. It was investigated whether a two-step approach would be more suited to the purpose of this study as this approach is advantageous if there is limited quantity of sample

and maximal optimisation of both steps is required (Takara, 2018). The High-Capacity RNA-to-cDNA™ kit (Applied Biosystems) was used to convert RNA samples to cDNA samples of 100 ng/μL which were then stored at -20 °C until use. PCR was then attempted with the One-Step master mix previously described without addition of the RT enzyme as it was speculated that the results using this would indicate whether the RT reaction was the limiting stage. The housekeeping gene *gyrA* was used to set up a dilution series of the cDNA, 50 ng of neat sample was used and then dilutions of this of 1/10, 1/100 and 1/1000 were made. This dilution series did not indicate the same inhibition seen with the one step reaction. Precision was improved between replicates however the Ct values were slightly higher than those seen in the one-step reaction. As previously mentioned the maximum sample concentration advised for this master mix is 100 ng RNA. Assuming full translation into cDNA occurs it can be assumed that 100 ng cDNA is the maximum allowed. Primer concentrations were originally 200 nM. An assay was set up using 10 or 100 ng cDNA template and both 200 nM primers or 500 nM primers. Results indicated that using both 100 ng cDNA and 500 nM primer concentration was preferable to any other combination for precision and Ct, however the Ct values still fell within the range of 29 to 36 which was still high, although the negative controls indicated that this was amplification of the target. As previously mentioned all target genes including *gyrA* are single copy genes. To attempt to demonstrate this primers for 16S rRNA were used (Velineni and Timoney, 2015) as multiple copies of this gene are present in the *S. equi* genome. It was expected that the Ct values for this target would be lower, indicating that the later values for the genes of interest were due to the low copy number. The 16S rRNA primers were used in a PCR assay alongside *gyrA*, *hasA* and both *murA* paralogs. The Ct values were lower for 16s rRNA than the other target genes although not by as much as expected, however the precision was also greatly improved compared to the target genes which suggests that both copy number and primer efficiency were issues. The annealing temperature was adjusted from 59 °C to 60 °C as amplification of the *gyrA* primers

which had the highest T_m was variable and the annealing time was lowered from 30 seconds to 20 seconds to reduce the chance of primer dimers which were observed in some of the replicates. At this point the stock of the one-step master mix was exhausted and an alternative was sourced. The PowerUpTM SYBRTM Green Master Mix (Applied Biosystems) was obtained and used for subsequent PCR optimisation. An optimised reaction using one master mix will not always be successful with a new one and so changing the master mix requires optimal conditions to be re-established (Bustin and Huggett, 2017). The new master mix did not appear to be as compatible with conditions used prior to it. Previously *hasA* expression was not detected in the non-mucoid phenotype samples, however unexpectedly this changed and amplification was observed in some of the replicates. This was speculated to be a result of contamination of the sample. As the RNA samples which previously appeared to contain some form of solvent contamination were omitted from analyses an alternative sample for this phenotype was not readily available. Purification of the previously omitted samples was attempted through re-precipitation and thorough drying of the pellet prior to reconstitution. Use of a different sample confirmed the presence of contamination in the original however precision and high Ct values continued to be a problem in this phenotype which possibly indicated inhibiting contamination was still present within the samples. Fresh bacteria pellets were prepared and RNA extracted as described previously. Fresh RNA samples were measured by NanoDropTM 2000 and were assessed using a 2% agarose gel as described previously. Unexpectedly the samples were found to be degraded following extraction. This had not been a problem prior to this and was hence attributed to contamination within some of the reagents used to precipitate RNA following DNase treatment. However upon performing another extraction omitting the DNase treatment step the samples were degraded again. Prior to conducting further troubleshooting the quantity of master mix left was assessed and it was concluded that there was not a sufficient amount to allow the optimisation of the assay and the obtaining of reliable results.

Unfortunately budget constraints meant that more PCR master mix or reagents for RNA extraction could not be obtained.

In conclusion, the design of primers for the low GC content genes of *S. equi* is challenging and the primers utilised for this project were likely not optimal. In addition to this the low copy number of the target genes likely contributed to the low precision and high Ct values obtained. Beyond budget constraints it would be desirable to start from scratch in order to take the time to test variations in primer criteria and their effectiveness. Also rather than using a commercial master mix kit it would be preferable to attempt to design a bespoke master mix where the concentrations of components such as magnesium (Mg^{+2}) which can affect the efficiency of DNA polymerase enzyme may be adjusted as required and possibly the maximum template concentration may be higher than 100 ng. Different DNA polymerases are also more or less resistant to inhibitors and so the ability to find the appropriate enzyme for the purpose of this study would improve the likelihood of success.

PCR and RT-PCR has been utilised frequently for the identification of *S. equi* species in both diseased and asymptomatic horses (Webb *et al.*, 2013; North *et al.*, 2014; Boyle *et al.*, 2016; Pusterla *et al.*, 2018). They have also been employed for transcription analyses for a variety of genes including those involved in HA synthesis (Chen *et al.*, 2009b; Prasad, Ramachandran and Jayaraman, 2012; Steward, Robinson and Waller, 2016). Where possible (such as for *gyrA*), primer sequences previously detailed in literature were used under the premise that they were adequate due to their acceptance for publication, as details of their efficiency had not been described. It has been suggested that methods in studies utilizing RT-PCR and qPCR fail to be adequately reported in detail such that they might be useful, despite the publication of MIQE guidelines for reporting of PCR experiments (Bustin *et al.*, 2009; Bustin, 2017). The lack of information available on specific conditions such as primer efficiency and volumes of reaction components in studies involving the *S. equi* species reinforces this notion. Therefore whilst the genes involved in

HA production have previously been analysed using PCR methods, not enough information was available to assist in the development of the assay for this study. In addition, whilst PCR has been utilised to amplify both *murA* paralogs for mutagenesis and overexpression (Du *et al.*, 2000; Marcellin, Chen and Nielsen, 2010), the real time gene expression of the paralogs has not yet been investigated, possibly due to the complexity of specific primer design experienced within this study.

Despite the difficulties in obtaining gene expression results the work in this chapter was able to demonstrate that whilst the *S. equi* species can be classed as susceptible to Phosphomycin under the conditions tested it is possible to supplement the media with non-inhibitory concentrations in order to test the effects of the antibiotic on other metabolic pathways. Whilst supplementation of growth media with the antibiotic did not have the expected effects on HA production, it is possible this was due to the implementation of resistance mechanisms and further work is needed to examine this. Future work will be discussed in Chapter 6.

5. Targeted proteomic comparison of *S. equi* strains

5.1. Introduction

Subspecies of *Streptococcus equi* are responsible for severe instances of disease in domesticated animals. Whilst *S. equi* subsp. *equi* is restricted to horses, *S. equi* subsp. *zooepidemicus* is a commensal organism and opportunistic pathogen in a wide range of hosts and instances of human infection have occurred (Timoney, 2010). *S. equi* subsp. *equi* is the pathogen responsible for Strangles, the most commonly observed infectious disease in horses, which affects lymph nodes in the head and neck and is highly contagious (Waller, Paillot and Timoney, 2011). Following the resolution of symptoms the bacteria are capable of persisting within the guttural pouch of the animal for up to 6 weeks in 50% of cases without detection unless by an invasive method such as lavage (Harrington, Sutcliffe and Chanter, 2002). *S. equi* subsp. *zooepidemicus* is capable of causing Strangles in horses, however it is also responsible for causing a variety of diseases in an array of hosts, such as mastitis in sheep, goats and cattle and respiratory diseases like pneumonia in ruminants, pigs and dogs and nephritis, septicaemia and meningitis in humans whom it may be transmitted to through contaminated dairy products (Steward *et al.*, 2017).

Virulence factors of *S. equi* subsp. *equi* include the HA capsule, Streptokinase, Streptolysin S and antiphagocytic M protein (SeM). These are also present in *S. equi* subsp. *zooepidemicus* with the exception that an M-like protein (SzP) is present instead of SeM and the two proteins function slightly differently from each other. In addition, homologues of other surface proteins are present in both subspecies (Timoney, 2004). Both subspecies share around 80% genome homology with the human pathogen *S. pyogenes* (Holden *et al.*, 2009) and upwards of 96% genome homology with each other (Timoney, 2004). As such it has been speculated that *S. equi* subsp. *equi* and *S. equi* subsp. *zooepidemicus* evolved from a common ancestor (Waller, Paillot and Timoney, 2011).

In 2004, a previously unclassified gram positive cocci with >70% genetic identity with *S. equi* subsp. *equi* and *S. equi* subsp. *zooepidemicus* was isolated during a routine bacteriological study investigating causative agents of mastitis in small ruminants. Following phenotypic, biochemical and genetic analysis the organism was proposed to be a new species and designated *Streptococcus equi*

subsp. *ruminatorum* (Fernandez *et al.*, 2004). Following that publication the strain was attributed to Strangles-like disease affecting spotted hyenas and plains zebras in Tanzania, where several clones of the strain were identified throughout the wild population (Honer *et al.*, 2006). A study following these strains then attempted to establish differences and similarities of these wildlife strains to the originally isolated strain by Fernandez *et al.* (2004) as well as type strains of *S. equi* subsp. *equi* and *S. equi* subsp. *zooepidemicus* using biochemical and molecular biological methods. Previously in this study it was revealed through HPLC assessment of culture media that SER17037 is a competent producer of HA capsule and competes sufficiently with the currently used industrial strain in terms of production according to OD (Chapter 2, Table 8). Whilst further investigation into SER17037 fermentation is required, the results suggest it may be a suitable strain for industrial production of hyaluronic acid and further characterisation is warranted.

Protein composition (proteome) is a useful way to assess phenotype variations and differences between strains. Differences and similarities in chemical characterisation, 16S rRNA sequence and M-like protein sequence between *S. equi* subsp. *ruminatorum* and the other *S. equi* strains were reported previously (Speck *et al.*, 2008), however the comparison of protein profiles were not presented and little more has been published about this novel subspecies of *S. equi*. *S. equi* subsp. *ruminatorum* DSMZ strain 17037 was obtained as a comparison strain in this work. Notably, phenotypic variants have been observed in this work (Chapter 2) for both SE40327 and SER17037 and the protein profiles of variants were analysed to compare whole cell protein profiles between strains of *S. equi* as well as differences between capsulated and un-encapsulated variants. Both SEE20561 and SEZ20727 have previously behaved as non-producers of HA under the conditions used in this study.

The most established method for analysing the protein content of bacteria cells is SDS-PAGE gel, (Emerson *et al.*, 2008). The resulting profile is a series of bands, often each band corresponds to more than one type of protein of similar size. The method has been successfully used in the analyses of both surface and internal proteins of bacteria (including *Streptococci*) as well as being used as a tool for distinguishing between and within bacterial species (Sheriff and Mahendraku, 2015). Using this method to identify bacteria species has been

described as fast, easy and inexpensive, and there have been reports that results of SDS-page profiling correspond to results of DNA/DNA or 16S rRNA based hybridization assays (Singhal, Maurya and Viridi, 2019).

In this Chapter the whole cell protein profiles of the strains described in Chapter 2 were compared along with an additional mucoid strain obtained from Hyaltech Ltd designated WS2/37. As previously mentioned, two phenotypes have been observed for SE40327 and SER17037, these were designated mucoid variants (MV) and non-mucoid variants (NMV). The phenotypes were separated and repeatedly passaged on agar to determine whether reversion to the alternate phenotypes occurred (data not shown). Agar was also switched between sulphur limited, 5% horse blood and BHI media. No phenotype switching was observed following up to 9 passages. The whole cell protein profiles of both types were separately compared for each of these strains.

5.2. Materials and Methods

5.2.1. Protein extraction

Bacteria were grown on sulphur limited agar plates at 37°C for 48 hours. Single colonies were inoculated into 50 mL of 37°C BHI media and incubated over night with 150 RPM shaking. The optical density at 660 nm was measured and each culture was corrected to 0.5-0.6 in 37°C BHI media. For the mucoid strains, 1x volume of 0.2 mg/mL hyaluronidase in BHI was added and the cultures incubated for a further 30 minutes to remove capsule. The cultures were then centrifuged at 4969 x g for 10 minutes. Pellets were washed twice in 500 µL PBS with centrifugation at 4969 x g in between washes. Pellets were then resuspended in 200 µL lysis buffer (60mM Tris, pH 8; 10mM MgCl₂; 50µM CaCl₂; 20µL/mL DNase and RNase) and 10 µL (250 U) of Mutanolysin (Sigma-Aldrich) was added. Following 1 hour incubation at 37°C the samples were placed in - 80°C for 1 hour. Samples were thawed at 37°C and a further 4 µL of DNase and RNase was added and incubated at 37°C for 30 minutes. SDS (20 µL) was added to a final concentration of 0.5 % and the samples were placed in a water bath at 90 °C for 30 minutes. Remaining debris was removed by centrifugation at 4969 x g for 5 minutes and the supernatant was transferred to a fresh 1.5 mL Eppendorf tube. The same volume of

Tris-Glycine SDS sample buffer (2X) (Novex™) was added and the samples were boiled for 2 minutes before being stored at -20°C until use. Samples were run on 12% Tris-Glycine Plus Midi gels (Invitrogen) at 175 V for 1 hour and 30 minutes. A PageRuler™ unstained protein ladder (ThermoFisher) was used, 5 µL of ladder and 10 µL of protein sample was used for each well. Completed gels were washed three times for 5 minutes in deionised water and stained with SimplyBlue™ SafeStain (ThermoFisher) overnight. The stain was then removed and the gels washed in deionised water until the background was clear and the bands were visible. Gels were imaged using a Gel Doc EZ Imaging System and photographs taken with Image Lab Software (Biorad).

5.2.2. Proteomic analysis

The selected strains from the whole protein profile gels were SEZ20727 and SER17037 both mucoid and non-mucoid phenotypes. These were run on a separate gel under the same conditions but with empty lanes separating them as per figure 42 to minimise potential sample spillover. The gel samples were then submitted to the Moredun Proteomic Facility (MPF) for analysis of the highlighted bands by ion trap liquid chromatography-mass spectrometry (LC-MS/MS) using an electrospray ionisation (ESI) interface. This was carried out using methods detailed previously (Hamilton *et al.*, 2018). The gel bands of interest were isolated by cutting from the gel before being destained, reduced, alkylated and digested using trypsin as per the procedure used by Shevchenko *et al.*, (1996). The digested protein samples were then transferred to low-protein-binding high-performance liquid chromatography (HPLC) sample vials. The LC-MS/MS-ESI was carried out using the Ultimate 3000 nano-HPLC system (Dionex) containing a WPS-3000 well-plate micro auto sampler, a UVD-3000 UV detector, a FLM-3000 flow manager and column compartment, an LPG-3600 dual-gradient micro-pump and an SRD-3600 solvent rack controlled by the chromatography software Chromeleon™ (Dionex). Sample volumes of 4 µL were directly injected. The final flow rate was maintained at 3 µL/min-1 through a 5 cm × 200 µm ID (polystyrene-divinylbenzene) monolithic reversed phase column (Dionex) sustained at 50 °C. Peptides were eluted using a 15 min linear gradient

from 8–45% solvent B (80% acetonitrile, 0.1% (v/v) formic acid) and were conducted through a 3 nl UV detector flow cell. For the process of ESI an interface of a 3-D high capacity ion trap mass spectrometer (amaZon-ETD, Bruker Daltonics) with a low-volume stainless steel nebuliser (50 µl/min-1 max) (Agilent). Tandem MS parameters utilised in this study are previously described (Batycka *et al.*, 2006). Proteins were identified using both the NCBI eubacteria database and the RAST annotated genomes of the strains of interest (See Chapter 3). Protein identifications were accepted if a minimum of two peptides were successfully identified and these peptides were represented by a minimum of four amino acid residues as an unbroken series of either *b* or *y* ions.

5.3. Results

5.3.1. Whole cell protein profile of *S. equi* strains

Figure 46 displays the whole cell protein profiles for the strains of interest in this study (SE40327, SEE20561, SEZ20727 and SER17037), labelled by their respective numbers, as well as mucoid (MV) and non-mucoid (NMV) variants of two of the strains (SE40327 and SER17037). Strains 40327, 40327MV, 40327NMV and WS3/37 are all working stocks of the same strain used in the industrial production of HA. From the figure it is observed that 40327 has a highly similar protein profile to 20561 and no obvious differences exist between 40327, 40327MV, 40327NMV and WS2/37. The profile of 20727 while also similar to 40327 appears to be slightly divergent in some areas, with the most obvious difference being the presence of bands higher up in the profile 85-100 kDa than for 40327 indicating the presence of some larger proteins, and a few areas of increased band thickness and clustering. Strain 17037 appears to possess slight variations in profile from both *S. equi* subsp. *equi* and *S. equi* subsp. *zooepidemicus* strains. Most notably there is the presence of two very thick, intense bands between 10-15 kDa size mark as well as a greater number of bands present below 10 kDa than in the other strains. No obvious differences in banding pattern are observed between 40327 mucoid and non-mucoid phenotypes. However the mucoid and non-mucoid variants of 17037 displayed some differences in profiles, most obviously the presence of a band in the mucoid variant about midway between

markers for 25 and 40 kDa which is not mirrored in the non-mucoid variant but appears to also be present at that approximate position in SEZ20727.

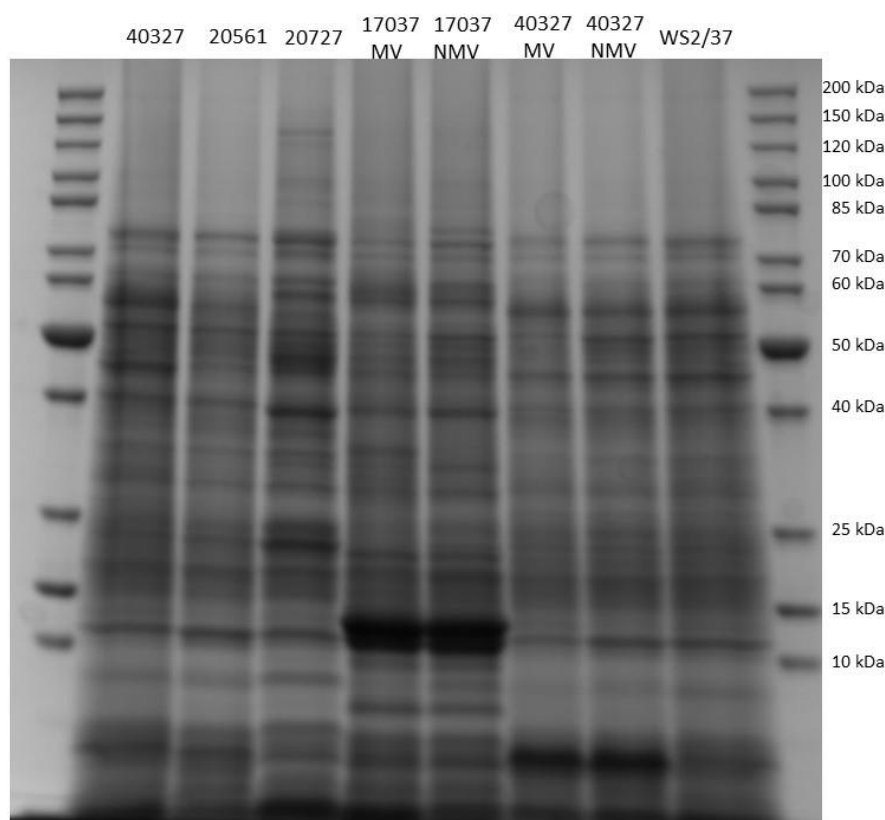


Figure 46 Whole cell protein profiles of *S. equi* strains of interest. Strains investigated were SE40327, SEE20561, SEZ20727, both mucoid (MV) and non-mucoid (NMV) phenotypes of SER17037 and SE40327 as well as a working stock which was claimed to have produced HA of 4.2 MDa (WS2/37).

5.3.2. Protein identification in bands of interest

As a result of the whole cell protein profiling of the strains a few bands of interest were noted and targeted for further investigation. The first was the band present in the mucoid phenotype of SER17037 but not in the non-mucoid phenotype of this strain. Secondly the distinctly intense bands at around 10-15 kDa size mark in both phenotypes of SER17037 which were not evident in any of the other strains. Hence these bands were selected for protein identification. The equivalent bands on SEZ20727 were also selected for analyses in order to compare the two strains. The gel submitted to Moredun Proteomic facility is displayed in figure 47 and the

bands isolated for protein identification are highlighted and assigned identifications.

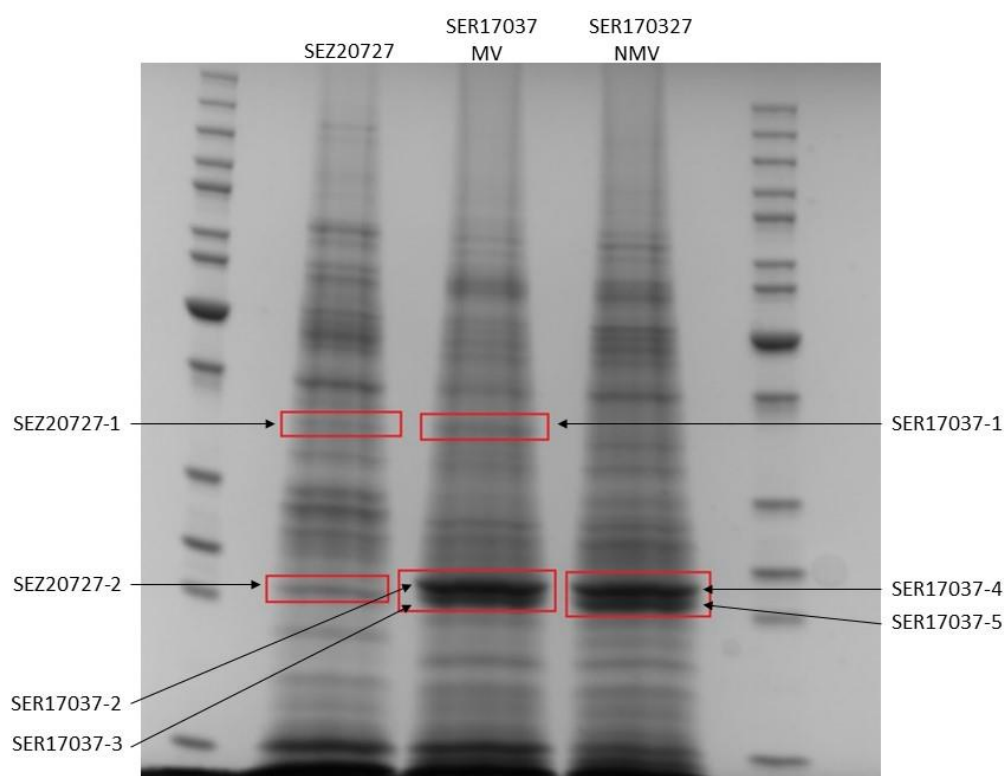


Figure 47 Protein gel of SEZ20727 and SER17037 mucoid and non-mucoid strains for proteomics analyses. Samples were run with single lane spaces between them to facilitate band isolation and extraction. The bands of interest isolated from the gel are indicated by the red boxes and arrows.

Table 13 and 14 display the proteins identified by both NCBI and RAST for the corresponding bands SEZ20727-1 and SER17037-1. Foldase PrsA was identified in both bands using both databases. SEZ20727-1 was also established to contain ATP-dependent 6-phosphofructokinase, a ribonucleotide reductase and a dipeptidase by NCBI and in addition the RAST annotation identified an enolase and a branched-chain amino acid aminotransferase. In SER17037-1 a peptidase and a hypothetical protein was identified using both databases and a peptidase identified using NCBI. Meanwhile the RAST annotated genome identified a protein matching to peg1459 was which corresponds to an antiphagocytic M protein.

SEZ20727-1			
RAST identification	MOWSE score	% Cover	No. peptides
6-phosphofructokinase	968.70	68.50	8
Foldase protein prsA 1 precursor	224.37	25.20	3
Ribonucleotide reductase of class Ib, beta subunit	215.84	16.90	3
Branched-chain amino acid aminotransferase	160.42	25.30	2
Proline dipeptidase	84.23	6.60	2
Enolase	64.07	8.50	2

Table 13 Protein identifications for band SEZ20727-1. This table displays the list of proteins identified from the RAST annotation of the SEZ20727 genome sequence as determined using Mascot 2.5.1 with MOWSE scores, % sequence cover and number of peptides identified.

SER17037-1			
RAST identification	MOWSE score	% Cover	No. peptides
Antiphagocytic M protein	755.72	22.80	10
FIG01118337: hypothetical protein	391.49	31.70	4
Foldase protein prsA 1 precursor	126.31	13.20	2

Table 14 Protein identifications for band SER17037-1. This table displays the list of proteins identified from the RAST annotation of the SER17037 genome sequence as determined using Mascot 2.5.1 with MOWSE scores, % sequence cover and number of peptides identified.

A larger number of proteins were identified in SEZ20727-2 (Table 15) compared to SEZ20727-1. Along with two hypothetical proteins and two proteins assigned as elongation factor Tu, NCBI Eubacteria database identified superoxide dismutase, Glx24 family stress protein, ribosomal protein L4 and a Uracil phosphoribosyltransferase. These four proteins were also identified in the RAST annotation as well as a ribosome recycling factor, alkyl hydroperoxide reductase protein C, an elongation factor Tu and a predicted 4-deoxy-L-threo-5-hexosulose-uronate ketol-isomerase. Superoxide dismutase was also identified in SER17037-3 and SER17037-5 as well as partial sequence being identified in SER17037-2. Additionally for SER17037-2 an YSIRK-type signal peptide containing-protein and another hypothetical protein were identified using the NCBI Eubacteria database whereas RAST only identified a hypothetical protein and antiphagocytic M protein. SER17037-2 and SER17037-4 are corresponding bands between mucoid and non-mucoid phenotypes of SER17037. The proteins identified in SER17037-4 were YSIRK-type signal peptide containing-protein, ribosomal protein s4 and a hypothetical protein as well as antiphagocytic M protein identified using the RAST

annotated genome. This corresponds partially to the identifications of the proteins in SER17037-2.

SER17037-3 and SER17037-5 are also corresponding bands on the protein gel between phenotypes. The following proteins were identified in both bands. Superoxide dismutase, antiphagocytic M protein, YSIRK-type signal peptide containing-protein, ribosome recycling factor and general stress proteins of the Gls24 family. Additionally identified in SER17037-5 were ribosomal protein L4, elongation factor Tu and hypoxanthine phosphoribosyltransferase. Both bands also contained a hypothetical protein.

SEZ20727-2			
RAST identification	MOWSE score	% Cover	No. peptides
Manganese superoxide dismutase	643.69	68.70	9
General stress protein, Gls24 family	458.90	51.40	6
Ribosome recycling factor	372.10	49.20	5
predicted 4-deoxy-L-threo-5-hexosulose-uronate ketol-isomerase	364.04	34.30	5
LSU ribosomal protein L4p (L1e)	327.41	41.10	4
Uracil phosphoribosyltransferase	183.99	21.40	4
Translation elongation factor Tu	124.96	5.80	2

Table 15 Protein identifications for band SEZ20727-2. This table displays the list of proteins identified from the RAST annotation of the SEZ20727 genome sequence as determined using Mascot 2.5.1 with MOWSE scores, % sequence cover and number of peptides identified.

SER17037-2			
RAST identification	MOWSE score	% Cover	No. peptides
Antiphagocytic M protein	1007.05	24.20	5
FIG01118337: hypothetical protein	207.52	15.30	2

Table 16 Protein identifications for band SER17037-2. This table displays the list of proteins identified from the RAST annotation of the SER17037 genome sequence as determined using Mascot 2.5.1 with MOWSE scores, % sequence cover and number of peptides identified.

SER17037-3			
RAST identification	MOWSE score	% Cover	No. peptides
Antiphagocytic M protein	962.99	22.80	6
Manganese superoxide dismutase	777.26	67.70	10
Ribosome recycling factor	364.73	43.80	2
General stress protein, Gls24 family	190.26	27.90	3
FIG01118337: hypothetical protein	108.66	6.90	2

Table 17 Protein identifications for band SER17037-3. This table displays the list of proteins identified from the RAST annotation of the SER17037 genome sequence as determined using Mascot 2.5.1 with MOWSE scores, % sequence cover and number of peptides identified.

SER17037-4			
RAST identification	MOWSE score	% Cover	No. peptides
Antiphagocytic M protein	892.43	22.40	7
SSU ribosomal protein S4p (S9e)	328.59	33.50	4
FIG0118337: hypothetical protein	81.11	8.70	2

Table 18 Protein identifications for band SER17037-4. This table displays the list of proteins identified from the RAST annotation of the SER17037 genome sequence as determined using Mascot 2.5.1.

SER17037-5			
RAST identification	MOWSE score	% Identity	No. peptides
Manganese superoxide dismutase	1019.76	82.10	8
Antiphagocytic M protein	831.16	22.40	7
LSU ribosomal protein L4p (L1e)	428.12	41.10	6
General stress protein, Gls24 family	358.41	32.40	5
Ribosome recycling factor	320.69	39.50	4
Translation elongation factor Tu	140.94	13.80	3
Hypoxanthine-guanine phosphoribosyltransferase	89.56	11.10	2

Table 19 Protein identifications for band SER17037-5. This table displays the list of proteins identified from the RAST annotation of the SER17037 genome sequence as determined using Mascot 2.5.1.

5.4. Discussion

Whole cell protein profiling is an approach which can be used to distinguish within bacteria species or strains. The most established method involves running extracted proteins on an SDS-PAGE gel which separates them based on molecular weight and produces a pattern of bands as seen in fig 41 (Emerson *et al.*, 2008). As previously mentioned, *S. equi* subsp. *ruminatorum* is a novel strain which was originally isolated in 2004 from mastitis found in small ruminants (Fernandez *et al.*, 2004). Since then it has not been mentioned often in literature. A study in 2008 which isolated and identified this strain in hyenas and zebras performed some characterization of this strain, however not all results were reported in full, including the results of whole cell protein profiling (Speck *et al.*, 2008). For the purpose of this study *S. equi* subsp. *ruminatorum* CECT 5772 was acquired from DSMZ (DSMZ no. 17037) as a comparison strain to SE40327 as it is reportedly mucoid in phenotype. Upon receipt and culture of this strain it was observed that two phenotypes were present (2.3.1 figure 6). The strains displaying both phenotypes were separated and passaged

multiple times on a variation of sulphur limited agar, brain heart infusion agar and horse blood agar, however no reversion was seen in either case. As previously mentioned, mucoid and non-mucoid phenotypes have also been observed for SE40327. Comparison of these two variants formed part of this preliminary proteomic investigation.

5.4.1. Protein profiles of *S. equi* strains

Since two strains displayed more than one phenotype, both mucoid and non-mucoid phenotypes were compared as well as comparing the protein profiles of SE40327 to SEE20561, SEZ20727 and SER17037. WS2/37 was a working stock donated from Hyaltech which had produced HA of 4.2 MDa on two separate occasions during production. It was hypothesised that SE40327 would produce a migration band pattern most similar to *S. equi* subsp. *equi* strain and that SEE20561, SEZ20727 and SER17037 would all produce different banding patterns. This was the first time protein profiling data had been presented comparing *S. equi* subsp. *ruminantium* to other *S. equi* subspecies. All strains were similar in banding pattern, with SE40327 most closely matching SEE20561 as expected. Whilst SEZ20727 and SER17037 were similar to SE40327 and SEE20561 there were subtle differences such as prominent bands at smaller molecular weights (<80 kDa) in SEZ20727 which were not present in *S. equi* subsp. *equi* strains. SER17037 presented banding patterns which were different to both *S. equi* subsp. *equi* strains and SEZ20727. This result was considered to support conclusion that SE40327 is a strain of *S. equi* subsp. *equi* when considered alongside the results of the genome sequencing previously described. One of the stand out features of the migration pattern of SER17037 was the presence of particularly thick and intense bands around 10-15 kDa in size which were not nearly as prominent in any of the other strains. Whilst no obvious differences between mucoid and non-mucoid phenotypes in SE40327 were observed, in SER17037 there was a band observed 40-45 kDa in the mucoid phenotype which did not appear to be present in the non-mucoid. This band coincided with a band in the SEZ20727 profile indicating similar sized proteins within this strain. As a result of these observations the bands described were submitted for proteomic analysis using LC-MS/MS.

5.4.2. Protein identification in selected SDS-page bands

The use of ESI-MS for protein identification increases the accuracy over other methods such as matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI-TOF-MS) because it is possible to perform more than one fragmentation on the sample, allowing for a peptide fingerprint to be determined and compared to a database of proteins (Emerson *et al.*, 2008). Moredun Proteomics Facility has provided proteomics support for many similar investigations of bacteria (Bannoehr *et al.*, 2011; Watson *et al.*, 2014a; Watson *et al.*, 2014b; Shahin *et al.*, 2018). As indicated in these publications and previously explained, protein identifications were only considered accurate if a minimum of two peptides were identified by a series of four or more amino acid residues represented by as an unbroken series of either *b* or *y* ions. The search engine used was Mascot 2.5.1 (Matrix Science) which is a software for identification, characterisation and quantitation of proteins using mass spectrometry data. It was used to search NCBI Eubacteria database as well as the RAST annotations of the sequenced genomes. In some cases the proteins identified were different between databases. As the RAST annotation is specific to the genomes of the strains investigated this was regarded as the more accurate tool for protein identification from the amino acid sequences determined by ESI-MS, with the NCBI identifications provided as supplementary material (Appendix C).

5.4.2.1. Foldase PrsA

A protein which was identified in both the band from the mucoid strain of SER17037 (SER17037MV) and the corresponding band in SEZ20727 was foldase PrsA precursor. This protein is found ubiquitously in all gram-positive bacteria and is responsible for folding exported proteins into their final functional configurations (Scott and Barnett, 2006). Whilst there is not an abundance of literature about PrsA in *Streptococcus* it has been investigated briefly in *S. mutans* where it was found that strains deficient in this protein displayed diminished early biofilm formation as well as indications that PrsA deficiency affected glucosyltransferases (Guo *et al.*, 2012). Whilst it is accepted that protein folding is crucial for correct

localisation of membrane proteins and for the successful export of fully functioning external proteins, the mechanisms for processing the surface proteins of Streptococci are not entirely understood (Crowley and Brady, 2016) and so the relevance of the presence of foldase PrsA precursor in the 17037MV and SEZ20727 bands and subsequent absence of this band in 170327NMV is uncertain. The few studies in *S. mutans* referenced have produced knockout mutants of PrsA and assessed their impact on growth and morphology as well as the effect on functionality of excreted proteins and biofilm formation. As the PrsA-deficient mutant of *S. mutans* developed by Guo *et al.* (2012) presented reduced synthesis of glucan, which was thought to be caused by impeded translocation of glucosyltransferase enzymes involved in production of glucan and other extracellular polysaccharides, the development of a knockout mutant of PrsA in *S. equi* would help determine whether this protein may also play a role in hyaluronic acid production through modulation of the enzymes involved either directly with its synthesis or in competing pathways. Hyaluronan synthase is a membrane integrated glycosyltransferase (Bi *et al.*, 2015) but the process of its integration into the membrane upon synthesis has not previously been revealed. Foldases such as PrsA are responsible for the maturation and folding of proteins destined for the cell membrane (Jakob *et al.*, 2015), so it could be speculated that foldases may play a role in the membrane integration of the hyaluronan synthase enzyme.

5.4.2.2. Antiphagocytic M-like protein

In addition, SER17037-1 was identified as containing Antiphagocytic M-protein. The molecular weights of M-like protein in *S. equi* subsp. *zooepidemicus* and *S. equi* subsp. *equi* were previously found to be 39 and 59 kDa respectively (Timoney, Artiushin and Boschwitz, 1997), however the SER17037-1 band is below 40 kDa in size, indicating that either the M-like protein of *S. equi* subsp. *ruminatorum* is smaller than that of other *S. equi* species or that this band contains a fragmented molecule. The identified sequence in SER17037-1 did not appear to align well with the sequence previously characterised as

M-like protein in this strain (Appendix D) (Speck *et al.*, 2008), however the sequence previously characterised is only partial and as discussed below divergence of the sequences of this protein is not uncommon. A BLAST of the RAST annotated sequence in NCBI resulted in a 100% match of 100% query cover to a protein in *S. equi* subsp. *ruminantium* CECT 5772 annotated as “streptococcal protective antigen” (accession number KED04751.1). Additionally the sequence matched 100% to 90.44% of the query to a sequence for M-like protein in *S. equi* subsp. *zooepidemicus* (accession number AHA95473.1). The M protein has been found to vary significantly in size between serotypes and even between strains within serotypes, an occurrence which is speculated to be due to the large number of repeats present within the gene (Fischetti, 2016). Also the hypervariability of the M protein sequence has in the past allowed for differentiation between strains of *S. pyogenes* (Metzgar and Zampolli, 2011), suggesting that significant variation between subspecies is not an unusual occurrence. A comparison of the proteins sequences of genes annotated in RAST as M-like proteins for strains of *S. equi* species is presented in Figure 48.

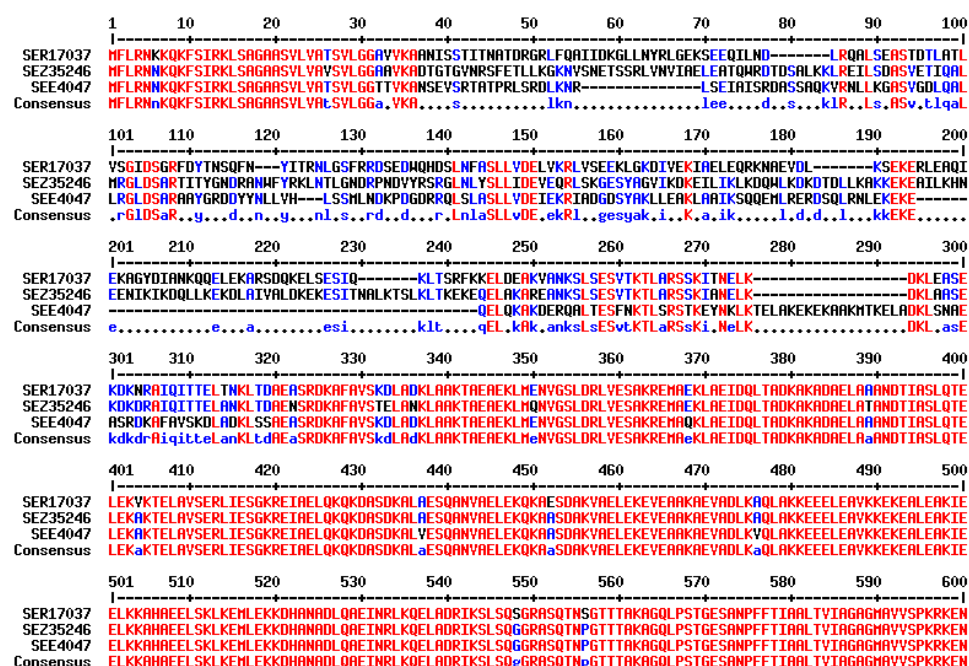


Figure 48 Protein sequence comparison of M-like proteins of *S. equi* species. Protein sequences of *S. equi* subsp. *equi* strain 4047 (SEE4047), *S. equi* subsp. *zooepidemicus* ATCC 35246 (SEZ35246) and *S. equi* subsp. *ruminatorum* (SER17037) were identified using RAST annotated genomes and compared using Multalin online software. The position of the LPST motif is indicated by underlining at positions 566-569.

As previously mentioned the data from whole cell protein profiling of this strain was described but not presented, therefore this is the first demonstration of the protein profile of *S. equi* subsp. *ruminatorum*. The high intensity of the bands observed in both phenotypes of SER17037 at 10-15 kDa prompted investigation by ESI-MS as this was distinctive to this strain. SER17037-2 and SER17037-4 were corresponding bands between mucoid and non-mucoid phenotypes of SER17037 respectively. Antiphagocytic M protein was identified in both of these bands, as was a protein which was confirmed as hypothetical in both the RAST annotation and NCBI eubacteria database. Once again the molecular weight of this band is far below that expected for Antiphagocytic M protein which possibly suggests fragmentation or processing. SER17037-3 and SER17037-5 correspond to the lower intense bands in the mucoid and non-mucoid phenotypes respectively and antiphagocytic M protein was also

identified in these bands but not in the corresponding sized band in SEZ20727-2.

The M protein is a surface protein identified in Group A *Streptococcus* (GAS) such as *S. pyogenes* and has been widely studied due to its established role as one of the main virulence factors (Smeesters, McMillan and Sriprakash, 2010). The GAS M protein covers the surface of the cell and provides protection from innate immune responses. This protective ability is further enhanced by the hypervariable region located at the N terminus of the protein which is a target for antibodies but varies in sequence between serotypes (Ghosh, 2018). It also has a role in colonisation and adhesion during the establishment of infection within the host (Brouwer *et al.*, 2016). *S. equi* subsp. *equi* and *S. equi* subsp. *zooepidemicus* also produce proteins which show homology to the M protein of *S. pyogenes*, these are designated M-like proteins (Harrington, Sutcliffe and Chanter, 2002). Similarly to *S. pyogenes*, these proteins have been shown to play a role in virulence by means of augmenting evasion of host immune system and adherence and colonisation of the host (Libardoni *et al.*, 2013). Whilst the sequence detected in the protein bands of interest in SER17037 were identified by RAST as Antiphagocytic M protein, a number of other genome sequences were annotated as M or M-like proteins in this strain. These sequences have been annotated in NCBI as various proteins such as protective antigen-like protein, *Emm*-like cell surface protein, antiphagocytic M-like protein, YSIRK-type signal peptide containing protein, cell surface-anchored protein and fibrinogen-binding cell surface anchored protein highlighting inconsistency in annotations. The sequence identified in the SER17037 protein band and the sequence which was RAST annotated as antiphagocytic M-like protein were searched in NCBI Blast with the search constrained to *S. pyogenes* and interestingly the former identified more closely with an *S. pyogenes* M protein at 37.92% identity of 46% query cover compared to 29.32% identity of 29% query cover. This indicates that although the latter sequence was annotated as an antiphagocytic M-like protein and the protein band

sequence as a protective streptococcal antigen of *S. equi* subsp. *ruminantium* in NCBI, the streptococcal antigen protein is actually closer in sequence to the M protein of *S. pyogenes*. Whilst the genome of *S. pyogenes* strain TJ11-001 contains only one gene designated as M-protein, the genomes of all the *S. equi* strains possessed many genes which are annotated as M-like proteins. Of the genes annotated this way in SER17037 none possess more than 25.65% identity to 98% of the *S. pyogenes* sequence. Similar results were seen when comparing the sequences from SEE4047 and SEZ35246 (data not shown). It is possible that not all of the genes annotated in the *S. equi* species are functional, however as it has been previously reported that *S. equi* subsp. *equi* possesses two types of M-like protein (Segura and Gottschalk, 2005), it is possible that other similar proteins exist with unknown functions. Although analysis of M-proteins was not an intended aim of this work, this has highlighted the need for a more systematic consideration of these significant surface proteins of Streptococci.

5.4.2.3. Superoxide dismutase

In addition to Antiphagocytic M protein, a Manganese superoxide dismutase was identified in SER17037-3 and SER17037-5. Superoxide dismutases (SODs) are enzymes involved in protecting cells from oxidative stress, they convert potentially damaging superoxide ions to hydrogen peroxide (H_2O_2) and molecular oxygen (O_2) and whilst bacteria produce many different types of SODs, only manganese superoxide dismutase (Mn-SOD) has been identified in *Streptococci* (Staerck et al., 2017). Evidence suggests that Mn-SOD is a virulence factor in group B *Streptococcus* species *S. agalactiae*, with deficient mutants exhibiting increased susceptibility to bacterial killing by macrophages in a mouse model of infection (Poyart *et al.*, 2001). Macrophages produce reactive oxygen species (ROS) and reactive nitrogen species (RNS) through a mechanism called oxidative burst in order to kill bacteria which they have engulfed into compartments named phagosomes and so the ability to neutralise reactive species may allow the bacteria to persist within the cells

(Slauch, 2011). The presence of this protein is presumably due to the fact the strains were grown aerobically with shaking. Previously a proteome analysis of *S. equi* subsp. *zooepidemicus* ATCC 35246 identified SOD in a spot of a 2-DE gel (Marcellin *et al.*, 2009) and proteomic analyses of other *Streptococcus* species have also identified SOD under standard growth conditions (Chaussee *et al.*, 2006; Wang *et al.*, 2012).

5.4.2.4. General stress proteins

Another notable protein identified in SER17037-3 and SER17037-5 was a General stress protein from the Gls24 family. The genome for SER17037 is available in NCBI (CECT5772) and this sequence is annotated as a hypothetical protein (KED05407.1). A second paralog of this gene is also present within the SER17037 genome which is also annotated as hypothetical protein (KED05405.1), the two paralogs possess 62.59% identity over 60% sequence cover. A BLAST of both sequences results in 100% identity to 100% query cover Asp23/Gls24 family envelope stress response protein of *S. equi*, indicating that the corresponding proteins in SER17037 could be reannotated. The definitive function(s) of Gls24 are still obscure. However its expression has been associated with multiple stresses including those caused by glucose starvation, exposure to bile salts and cadmium chloride (CdCl₂) as well as zinc and copper exposure and oxidative stress and additionally virulence in an endocarditis model in mice (Hartke *et al.*, 1998; Svensater, Sjogreen and Hamilton, 2000; Nannini *et al.*, 2005; Stoyanov *et al.*, 2010). Proteins within this family have an average homology of 60% and in addition they have been found to have an average of 30% homology to the alkaline stress protein 23 (Asp23) which was discovered in *Staphylococcus aureus* (Giard *et al.*, 2002). Thus this family of proteins has been characterized as the Asp23/Gls24 superfamily and it has been speculated that due to their position within operons they may have properties as transcriptional regulators (Ferrandiz *et al.*, 2019). Due to the uncertainty around the function of the Gls24 proteins the significance of their presence in the distinct protein bands of

SER17037 is unclear. The sequences identified in the RAST annotation as Gls24 family stress proteins were identified in the NCBI Eubacteria database as a hypothetical protein and a transcriptional regulator. However a BLAST of the sequence corresponding to the RAST peg number resulted in a 100% match to Asp23/Gls24 family envelope stress response protein of *S. equi* (accession WP_037578388.1). This protein was also identified in the corresponding band of SEZ20727-2, as the cultures were grown to stationary phase prior to protein extraction it is not unusual to see the presence of this protein as general stress proteins are upregulated during this phase of growth (Gottesman, 2019). It is possible the presence of this gene product is detected due to the bacteria being grown in batch culture and the proteins being extracted during stationary phase where glucose starvation may occur and pH decline may induce stress responses. Previously it has been suggested that some stress proteins are involved in responses to multiple stresses such as starvation and acid shock (Svensater, Sjogreen and Hamilton, 2000).

5.4.3. Major surface structures of *S. equi*

Although no obvious links to HA production were observed in the protein bands investigated in this chapter the proteins identified appeared to be primarily surface based and/or virulence factors.

5.4.3.1. M-like protein and superoxide dismutase

Both M protein and HA capsule have long been considered virulence factors in *S. equi* species (Timoney, 2004). Both have been found to be necessary for protection against phagocytosis by immune cells (Dale *et al.*, 1996b; Timoney *et al.*, 2014) and HA has been found to mediate the adherence of group A *Streptococcus* cells to human keratinocytes through binding of M protein to CD46 (Schrager, Rheinwald and Wessels, 1996). Both of these virulence factors have been found to contribute to phase switching to the hyperinvasive serotype of *S. pyogenes* designated M1T1 (Cole *et al.*, 2010) and *S. equi* isolated from asymptomatic horses displayed truncated M

proteins and as a result an increased susceptibility to phagocytosis, indicating that M protein plays a role in both colonization and virulence (Chanter *et al.*, 1999). The M protein is considered a stereotypical example of a surface protein which anchors to the cell wall via C-terminal LPXTG motifs (Fischetti, 2016).

Due to its previously described protective action against bacteria killing within the phagosome the superoxide dismutase (SOD) is also considered a virulence factor, with mutants deficient in the SodA gene suffering enhanced susceptibility to killing by macrophages, sensitivity to oxidative stress and decreased pathogenesis in a mouse model (Poyart *et al.*, 2001). As previously mentioned this protein could be present due to the conditions used to culture the bacteria.

5.4.3.2. Foldase PrsA

Although not considered in itself relevant to HA production, the presence of foldase PrsA may be significant due to its involvement in post-translational processing of proteins in the extracytoplasmic region of the cell (Scott and Barnett, 2006). PrsA is reportedly found in all gram positive bacteria and its functions have been attributed to biofilm formation, processing of stress response proteins including molecules involved in antibiotic resistance, assembly of virulence factors, integration of surface proteins into the cell envelope and even the production of polysaccharides such as glucan (Guo *et al.*, 2012; Crowley and Brady, 2016). Hyaluronan synthase is a progressive glycosyltransferase which is integrated into the membrane where it both synthesises and exports the HA chain (Bi *et al.*, 2015). HA has previously been considered a virulence factor (Wessels *et al.*, 1991; Stollerman and Dale, 2008) and whilst the mechanism of its production have been extensively studied, little information is available on the mechanisms surrounding the synthesis and integration of the HAS enzyme. Whether foldases such as PrsA play a role in the folding and cell membrane integration of HAS is unknown but could be speculated due to the observed presence in the band in the mucoid variant of SER17037 but not the non-mucoid. The approximate molecular weight of PrsA is 36 kDa (Jakob *et al.*,

2015) which corresponds to the position of the band in the protein gel from this study.

5.4.4. Conclusions

Despite numerous investigations of HA production by various *Streptococci*, the factors influencing HA production are not all clear. As previously mentioned it has been linked to virulence, but also environmental stressors such as glucose availability, oxidative stress, mineral availability and pH have been described as influential aspects (Chong *et al.*, 2005; Huang, Chen and Chen, 2006; Duan *et al.*, 2008; Liu *et al.*, 2008; Don and Shoparwe, 2010; Pires, Eguchi and Santana, 2010). Thus it is entirely possible that many factors influencing the production of HA in pathogenic bacteria have not been considered. Whilst the process of HA production is a good place to start, it would also be logical to investigate the role of the mechanisms involved in the production and placement of the enzyme responsible for HA production.

Whilst the work in this chapter did not provide obvious avenues to follow in the elucidation of the mechanism of HA production, it has introduced ideas and questions surrounding hyaluronan synthase production, processing and integration into the membrane. Whilst many streptococcal surface proteins are anchored to the cell wall at the C terminal using the LPXTG motif (Fischetti, 2016), the HAS enzyme possesses numerous domains (6 in total) which are associated with the cell membrane (Weigel and Baggenstoss, 2012). The post-translocational folding and integration of an enzyme to the membrane at so many domains is likely to be a delicate process and whether mis-integration of some of the domains could impair function may be speculated. It has previously been identified that HAS activity is regulated by the presence of cardiolipin molecules which are associated with the active enzyme (Tlapak-Simmons *et al.*, 1998) however the point at which this association occurs and how it is regulated is not known. It is possible that these are aspects concerning HA production which have not been investigated in full and further clarification of the processes surrounding optimal enzyme function would be beneficial.

6. Discussion and Conclusions

6.1. Fermentation of *S. equi* species for HA production

Since its identification in the bovine vitreous humour in 1934, HA has been produced through a variety of methods including extraction from human umbilical cords, animal sources such as the combs of roosters and most recently fermentation of naturally HA producing bacteria (Necas *et al.*, 2008). As a valuable product in the pharmaceutical, cosmetic and biomedical industries (Sze, Brownlie and Love, 2016) the development of optimal manufacturing procedures is an important area of research. This study aimed to evaluate HA production in a variety of *S. equi* strains, including a strain currently utilised by Hyaltech Ltd in the manufacture of hyaluronic acid based medical devices, *S. equi* NCIMB 40327 and the following strains from DSMZ, *S. equi* subsp. *equi* 20561, *S. equi* subsp. *zooepidemicus* 20727 and *S. equi* subsp. *ruminatorum* 17037.

6.1.1. Advantages and disadvantages over animal production

Whilst the biotechnological production of HA has predominantly taken over from animal extraction, there are a few caveats to it (Badri *et al.*, 2017) and the extraction of HA from animal sources still continues. Rooster combs contain a high concentration of high molecular weight HA, however due to ethical concerns, difficulties with the purification processes and an increased awareness of threat of cross species virus transmission and immune complications resulting from contamination with animal protein, the use of HA derived from animal sources is undesirable (Boeriu *et al.*, 2013). Bacterial HA production using pathogenic Streptococci has also received some criticism due to risk of contamination from pro-inflammatory molecules (Liu *et al.*, 2011). In addition, there are limitations to the fermentation process due to the high viscosity of the media upon hyaluronic acid secretion, meaning maximum concentrations are only 5-10 g/L and the average molecular weight of bacterial HA is often lower than that of animal derived (Blank, Hugenholtz and Nielsen, 2008). The polydispersity of HA produced by bacteria can pose a challenge for the manufacture of products containing the polysaccharide as the immunoreactivity of it is specific to the molecular weight, with low

molecular weight HA inducing pro-inflammatory responses and high molecular weight HA providing anti-inflammatory properties (Cyphert, Trempus and Garantziotis, 2015).

As such there are advantages and disadvantages to HA derived from both animals and bacteria. Most certainly in order to remain ethical the use of animals should cease. There appears to be promise in the area of biosynthetic HA production using non-pathogenic bacteria such as *B. subtilis*, *L. lactis* and *E. coli* engineered with the HA synthesis machinery from naturally producing strains (Yu and Stephanopoulos, 2008; Sheng *et al.*, 2009; Kaur and Jayaraman, 2016; Zhang *et al.*, 2016), however while HA is produced successfully the concentrations and molecular weights of the product are often greatly reduced compared to production in the indigenous species.

Therefore the fermentation of natural HA producers such as Group C *Streptococci* is at present considered the most adequate option for HA production for a variety of products. Many studies have been conducted on HA production in *Streptococcus* species concerning media type and culture conditions such as temperature, pH and dissolved oxygen content as well as other parameters such as shear stress and nutrient levels (Armstrong, Cooney and Johns, 1997; Duan *et al.*, 2008; Liu *et al.*, 2008; Velineni and Timoney, 2015; Zakeri and Rasaei, 2016), however many of these studies have focussed on an engineering approach. The fact that HA production likely did not develop in the bacteria being investigated for the purpose of human interest appears to have been largely ignored.

For this reason the present study aimed to attempt to return to the question of HA production by *Streptococci* from the bacteriological perspective in a bid to unravel the role of HA production in the complex system of pathogenic bacteria.

6.1.2. The fastidious nature of *Streptococcus*

Fastidious bacteria are those which have specific requirements for growth. Streptococci are facultative anaerobes which vary in nutritional requirements between strains due to the variation in the hosts targeted (Timoney, 2010). Both group A and C Streptococci are particularly

demanding with regards to organic nitrogen requirements, resulting in yeast extract, casein hydrolysate and peptones being common components for culture media for these bacteria alongside supplementation such as phosphate salts and magnesium (Armstrong, Cooney and Johns, 1997). Whilst these components are present in the media used in this study, occasions of absence of growth did occur and was likely due to reasons such as the media temperature when inoculated or the fitness of the cultures prior to frozen stock preparation. In the case of the original master stock of SE40327, the freeze dried culture was revived at NCIMB (<https://www.ncimb.com/>) where the collection was catalogued and delivered in two agar slants. Upon receipt of the slants growth was extremely minimal and only the growth from one of the slants was successfully transferred to and grown in liquid media before glycerol stocks were prepared. With regards to the other strains the bacteria were grown overnight in media for the preparation of glycerol stocks meaning they were likely in stationary phase of growth at time of freezing. Due to any number of the above mentioned conditions on some occasions frozen stocks were unsuccessfully revived on agar plates. On some occasions growth on agar was successful but upon colony transfer to liquid media growth was then absent following a period of culture time, an occurrence which is also experienced during the industrial process which can cause delays to production. Although not deemed a requirement at the outset, a recommendation would be the adoption of more standardised stock culture, storage and starter culture procedures to improve batch-to-batch consistency.

6.1.3. Current production procedure and its challenges

As previously discussed, fermentation of HA by Streptococci is most often carried out using a batch mode of fermentation (Armstrong and Johns, 1997; Liu *et al.*, 2011). However the process modelled in this study is carried out in continuous culture for reasons mentioned in chapter 2. Whilst in theory this is an appropriate approach, in practice it has been observed that HA production declines after a period of time regardless of maintenance of growth rate in continuous culture. As working stocks of the bacteria are frequently prepared from samples of culture from the

fermenter taken during a 'run', it is possible that this approach could affect the production during future runs, particularly if the decline has been caused by the development of mutations resulting in the non-mucoid phenotype such as observed in chapter 2. This poses the question as to whether continuing operating with a continuous culture is worthwhile or whether a batch culture system with options such as buffering would be more suitable and less costly than running a fermenter which is unreliable in producing HA. Another option could be running the fermenter for less time than the 20 days which is currently standard practice. The purification process is extensive with many filtration steps of the culture media prior to HA precipitation, due to the concerns mentioned above. Often this results in an approximately 30% yield of the initial concentration of HA in the culture media (personal communication, Hyaltech production staff - data not available). While not within the scope of this study, it is possible that an optimised purification procedure may allow for the use of batch fermentation in place of the current continuous method.

Another interesting observation of this study was that production did not appear to be significantly different between the currently used multicomponent fermentation media and off-the-shelf BHI media. This brings into question whether it is necessary in this case to use such a complex media for fermentation. Whilst previously it has been found that a high glucose concentration promotes greater molecular weight and higher concentration of product (Pires and Santana, 2010), this did not appear to be the case for production under the conditions used in this study. In addition to this the use of sulphur limited media to promote a mucoid phenotype was not supported in this study due to the observation that mucoid phenotype was enhanced on BHI agar compared to sulphur limited agar. In addition to greater mucoidy, bacteria inoculated onto BHI media required 24 hours to grow compared to a minimum of 48 on sulphur limited media, indicating that time could be saved in the preparation of inoculum if BHI media were used instead of sulphur limited and fermentation media.

6.1.4. *S. equi* subsp. *ruminatorum* as a candidate for HA production

The most commonly utilised native HA producing bacteria species are still *S. equi* subsp. *equi* and *S. equi* subsp. *zooepidemicus* (Kim *et al.*, 1996; Chen *et al.*, 2009a; Choi *et al.*, 2014). As it is a novel species, *S. equi* subsp. *ruminatorum* had not previously been investigated as a candidate strain for HA production. From the results of this study it would appear that this species is capable of producing similar concentrations of HA as the currently used strain. Although molecular weight was not specifically investigated due to technical difficulties, the similar retention times during HPLC could indicate similar molecular weights between the SER17037 and the currently used strain. More detailed characterisation of *S. equi* subsp. *ruminatorum* HA production through further studies might indicate whether this, too, could be useful in biomanufacture.

6.1.5. Future work

The results of this study suggest that changes to the current procedure could be appropriate following further investigation. The production and molecular weight of HA within BHI cultures compared to the current media on a pilot scale would provide more information on the possibilities of this media being used in its stead. The use of sulphur limited broth could possibly be omitted and the effects of this change assessed. In addition, further comparisons of SER17037 to the current strain are required, such as molecular weight of the HA product and whether there is an increased production of possibly contaminating molecules such as exotoxins compared to SE40327. These analyses would provide more information as to the suitability of this strain as a candidate for industrial production of HA.

6.2. Comparative genomics for insight into HA synthesis regulation

The sequencing of the genomes of the strains of interest in this study was carried out for a number of reasons. The genome of SE40327 had not previously been sequenced and this would allow updated characterization of the strain and support molecular experiments. In addition the ability to compare the genomes of the *S. equi* strains as well as compare them in practice provided multiple angles for investigations. The ability to compare the genome of *S. equi*

subsp. *ruminatorum* to the other species assisted in further characterisation of this novel strain which had not previously been carried out.

6.2.1. Challenges of DNA extraction of *S. equi* species

Gram positive bacteria such as Streptococci possess dense layers of peptidoglycan around their cell membrane (Silhavy, Kahne and Walker, 2010). Subsequently they have been found to possess greater resistance to cell lysis than gram negative bacteria, including resistance to lysis enzymes such as lysozyme (Chassy and Giuffrida, 1980). As a result the extraction of DNA from these bacteria can involve harsh reagents which are costly or required physical methods which, depending on what the DNA is required for (Moreira *et al.*, 2010), are more or less suitable.

Several methods were attempted in this study to extract DNA for genome sequencing. Initial attempts used lysozyme (1mg/mL) followed by extraction using the Qiagen DNeasy Blood and Tissue kit as this was readily available and widely used, however extracted DNA concentrations were too low for sequencing. Variations to the kit were attempted such as increasing the cell pellet used, however this blocked the spin columns. Incubation time with lysozyme was increased but did not have the desired effect. Sonication within the initial Qiagen lysis buffer prior to spin column processing was trialled but did not result in greater DNA quantity. In attempts to improve cell lysis, an SDS based lysis buffer was prepared containing 200mM Tris-HCl pH 7.5, 250mM NaCl, 25mM EDTA and 0.5% SDS. This has been used in a protocol for DNA extraction from plant cells and involved vortexing and incubating the cell pellet in the lysis buffer for up to one hour before centrifugation of sample and precipitation of the DNA in the supernatant using isopropanol and purification using 70% ethanol. Whilst greater concentrations of DNA were achieved they were not found to be pure enough to pass the initial quality control checks for sequencing carried out at Edinburgh Genomics.

The MasterPure™ DNA Purification Kit described in Chapter 3 methods had been successfully used for DNA extraction in *E. coli* within the lab. Following the challenges encountered with the Qiagen DNeasy Blood and Tissue kit, the Epicentre kit was sourced and extraction attempted with it.

While not perfect, this kit resulted in greater DNA concentration following extraction than had been previously achieved. A larger bacterial pellet was required for extraction than the protocol advised and so the volume of reagents were doubled. Thus the extraction of DNA from the *S. equi* strains was challenging and following the experiences of this study it may be stated that kit type and mechanical lysis techniques should be considered when extracting DNA from gram positive bacteria.

In addition to difficulties in DNA extraction another observation from this study is the over estimation of nucleic acids observed by the NanoDrop™ 2000. On several occasions the samples measured on this instrument appeared to be of adequate concentration and purity with regards to the criteria set out by Edinburgh Genomics. However upon submission of the samples they were found not to pass the criteria when analysed using EdGen quality control methods which include Qbit in measurements. The guidelines for NanoDrop™ 2000 do suggest that at low concentrations, readings may be inaccurate. This would suggest that whilst an efficient method for nucleic acid and protein quantification, the results of NanoDrop™ 2000 readings should perhaps be considered an overestimation.

6.2.2. Annotation, mis-annotation and assembly

The presence of contigs within the sequenced genomes of the strains of interest presented challenges with regards to evaluating the structure of the *has* operon due to the resulting lack of all genes being located and annotated. In all instances the *has* operons in the strains were split across two or more contigs, most often following *hasB*. The likelihood of this happening could have been increased by the duplication of *hasC* which following a comparison of the flanking sequences of the two paralogs is speculated to be the reason for the existence of the two copies (Blank, Hugenholtz and Nielsen, 2008). The presence of repeat sequences increases the likelihood that a contig will end at that gene as well as the likelihood of either one or both copies being lost (Treangen and Salzberg, 2011). As a result of this, in none of the strains sequenced was the whole of the *has* operon isolated to a single contig, meaning assessment of the structure could not be made as planned. In order to complete this the

whole genome of the strains would need to be resequenced in long read format which is outside of the resources of this study.

In addition to challenges posed by genome assembly, the variation between annotations from different databases and in different strains of the same sequences resulted in the time consuming task of confirming the identity of genes of interest. Whilst the use of multiple database annotations identify the true integrity of the RAST annotation is a start, there remains uncertainty, particularly as to whether the hyaluronidase genes identified are bacterial or phage genes. Until annotations between databases are coordinated this will continue to be a challenge in genomic and proteomic studies which require sequence confidence to be as certain as possible in absence of experimental evidence.

6.2.3. Selection of putative HA production-related targets

The genomes of mucoid and non-mucoid strains of SE40327 were compared in order to reveal potentially relevant genetic differences. A total of 47 genes were found to be less than 100% identical between the two phenotypes. For the purpose of this study genes contained within the same contig as the hyaluronic acid production machinery were the main focus. In addition to this the focus was concentrated around possible regulators of surface molecules and virulence factors due to previously mentioned associations between HA capsule and virulence, as well as the fact that the hyaluronan synthase enzyme is reported to be a surface membrane bound glycosyltransferase. However there were other genes found to be differentially regulated such as histidine protein kinases and transcriptional repressors which may be of interest however were deemed outwith the criteria for analysis in this study.

6.2.4. Future work

In order to confidently make comparisons and assumptions the whole genomes of the strains of bacteria should be sequenced should resource allow. An increased number of both HA producing and non-producing strains would make identification of targets for analysis to occur. In addition to this it would allow for the development of greater understanding of the structure of HA producing genetic machinery and its

regulators. It has been speculated by Blank, Hugenholtz and Nielsen (2008) that *hasD* and *hasE* are under control of more than one promoter and this is an avenue to be investigated further alongside the identification of the regulator of the second *hasC* paralog.

Further investigations into the genes identified as non-identical between mucoid and non-mucoid SE40327 may reveal other avenues of interest out with the general HA synthesis machinery, widening the scope for research into HA production regulation.

Comparative genomics is a useful tool for the identification of genetic targets for experimental investigation and can provide a copious amount of information to decipher and speculate over, allowing for more questions to be asked and hypotheses posed. Analyses of the genomes of the strains of interest should be continued alongside the design of experimental analyses.

6.3. The HA production pathway in competition with essential biosynthetic processes

Streptococcal HA production is facilitated through two pathways which produce the two precursors UDP-glucuronic acid (UDP-GlcA) and UDP-N-acetylglucosamine (UDP-GlcNAc) (Yamada and Kawasaki, 2005). However bacteria are efficient organisms and so a number of the products of the intermediate steps of these pathways are constituents for other cell wall products such as teichoic acids, antigenic polysaccharides and peptidoglycan, as well as being involved in other cellular processes such as the pentose phosphate pathway and glycolysis (Chong *et al.*, 2005). In the case of fermentation, lactic acid production also competes with HA synthesis for carbon source and increasing concentrations of lactic acid in cell culture inhibit both growth and HA production (Liu *et al.*, 2011). Therefore capsule production is consistently outweighed by other more essential processes and the ability to tip the balance with regards to emphasis on HA synthesis may result in elevated levels of production or a higher molecular weight product. Thus the regulation of the competing pathways of HA synthesis are an important area of research.

6.3.1. The MurA enzymes as a target

The identification of the relationship between increased molecular weight of HA and the downregulation of MurA by Marcellin, Chen and Nielsen (2009) highlighted the relevance of other pathways in the topic of HA synthesis by bacteria. As previously described, MurA competes directly with the hyaluronan synthase enzyme for the HA precursor GlcNAc. Considering the observation that GlcNAc is the limiting precursor with regards to molecular weight (Chen *et al.*, 2009b) it is logical that a reduction of MurA expression or activity could result in increased yield and/or molecular weight of HA due to greater availability of this precursor. Therefore the investigation into MurA and its function as an aspect of HA regulation is an important area for investigation.

As peptidoglycan is an essential component of the bacterial cell, interference with MurA is a delicate affair as complete inhibition of peptidoglycan synthesis will inhibit growth. Conveniently, gram-positive bacteria such as Streptococci possess more than one active copy of this enzyme and so the option exists to identify the most appropriate target out of two copies.

6.3.2. The use of phosphomycin for MurA inhibition

As previously mentioned in chapter 4, phosphomycin is a specific inhibitor of MurA. Whilst it had previously been claimed that it was inhibitory to Streptococci at concentrations of ng/mL this appeared to be implausibly potent and so investigating the MIC in the strains of interest was the reasonable first step. The observation that at lower concentrations in µg/mL phosphomycin does not inhibit growth of the strains of interest resulted in the opportunity to assess the influence of its presence on hyaluronic acid production. Whilst it was hypothesised that the presence of phosphomycin may result in increased production and molecular weight this did not appear to be the case upon observation of the results. This was postulated to be due to increased expression of *murA* as a mechanism for resistance, however this would need to be confirmed by additional analyses including gene expression analyses in the presence of phosphomycin. Due to challenges discussed below and in chapter 4 this remains to be

concluded. Furthermore, if phosphomycin had resulted in an increase in molecular weight of HA there could be challenges to introducing it into industrialised fermentation processes due to the fact it is an antibiotic and the issue of antibiotic resistance is an important topic in today's society. Whilst phosphomycin inhibition occurs through covalent binding of the antibiotic to the MurA enzyme (Silver, 2017) thus neutralising itself, the presence of free phosphomycin within fermentation waste may raise questions about the suitability of this method. This would need to be carefully investigated prior to establishing a fermentation method and a protocol for monitoring put in place.

6.3.3. Challenges due to paralog sequence similarity

Whilst the activities of the MurA paralogs in *S. pneumoniae* have been observed and found that MurA1 possess greater enzymatic activity than MurA2 (Du *et al.*, 2000) this has not been confirmed to be the case for *S. equi* species. On the assumption that the case is similar, the differences in expression of the *murA* genes has not been elucidated and it may be the case that one undergoes greater expression than the other. Due to the similarities of the paralog nucleotide sequences design of specific primers for RT-PCR were challenging and the presence of non-specific binding and double products which were apparent due to the observation of more than one peak in the melt curves, did not allow for confidence in results. Three rounds of primer design were attempted with only one pair found to be consistently specific for *murA1*. Due to the requirements for RT-PCR such as product lengths of no longer than 120 bases, primer lengths of approximately 20 bases, as few single base repeats as possible and preferably primers located within the middle of a gene the design of primers specific to the *murA* paralogs future attempts at primer design will not be straight forward.

6.3.4. Future work

The development of specific primers for expression analyses of both *murA* paralogs is important for future investigation into the differences between the two in terms of function. Following this task the effect of the presence

of phosphomycin in the culture media would be investigated in order to establish whether upregulation of the genes occur.

Whilst out with the available resources of this study it would be useful to develop a knockout mutant of each of the copies of MurA following confirmation that both are expressed. As it has been observed to be the less active copy, MurA2 would be the initial target as it is hypothesised the absence this copy would have a more minor effect on growth. Initiation of planning for an allelic exchange experiment was considered however a suitable plasmid could not be sourced within the resources of this study.

6.4. Proteomic tools for distinguishing *S. equi* species

Whole cell protein profiling for identification of bacteria types is a simple and well established method based on the assumption that different bacteria species and sometimes strains produce different banding patterns from each other and that the same species/strain will produce reproducible patterns under lab conditions (Singhal, Maurya and Viridi, 2019). The protein profiles of the novel species *S. equi* subsp. *ruminatorum* were reportedly compared to strains of *S. equi* subsp. *equi* and *S. equi* subsp. *zooepidemicus* (Speck *et al.*, 2008), however the resulting profiles were never presented by those authors. In addition to this, the observation of both mucoid and non-mucoid phenotypes present from stocks of the same strain prompted the protein profiles of the strains of interest in this study to be investigated.

6.4.1. Challenges of protein extraction of *S. equi* species

Extraction of proteins from *S. equi* subspecies was challenging for the same reasons described above for DNA extraction (6.2.1.). Cell disruption was attempted using sonication for physical disruption and lysis buffer with mutanolysin an enzyme discovered to be capable of lysis of Streptococci (Yokogawa *et al.*, 1974). In addition the MasterPure™ kit (Epicentre) was adapted to attempt to isolate proteins rather than DNA, however this was unsuccessful. The protocol ultimately used was adapted from a protocol for the extraction of lipopolysaccharides (LPS) from the cell wall of gram negative bacteria. Hyaluronidase was required prior to centrifugation as the presence of capsule inhibited the formation of a firm bacterial pellet at

less than 40,000 x g. The adapted protocol detailed in this study (chapter 5) would be suitable for other resilient bacteria species.

6.4.2. Proteomics for subspecies differentiation

From the results of SDS-PAGE analysis, variation in protein banding profiles were observed between species but were not notably different between strains of *S. equi* subsp. *equi*. Therefore it can be stated that protein profiling is suitable only for differentiation between *S. equi* subspecies and not strains of the same subspecies. While variation in protein profiles between phenotypes of SE40327 was not observed, the absence of a band in the non-mucoid phenotype of SER17037 which was present in the mucoid phenotype was noted and so was investigated further using proteomic analyses. The presence of variation between the phenotypes in only SER17037 suggests that it may not be applicable to other subspecies. However as this study has shown SER17037 appears initially promising as a HA production candidate and so any insight into HA production in this strain may be useful.

6.4.3. Evidence of M-like protein processing or degradation

The results of the proteomics analyses revealed Antiphagocytic M-like protein to be present in numerous bands within SER17037. The presence of this protein in multiple bands attributed to small sized proteins indicates the presence of fragmented proteins, presumably specifically cleaved by a protease. Whilst there have been many studies surrounding the function and presentation of the M proteins and M-like proteins (Smeesters, McMillan and Sriprakash, 2010; Metzgar and Zampolli, 2011; Brouwer *et al.*, 2016; Ghosh, 2018), little appears to be known about the processes involving M-protein turnover. It could be that in stationary phase degradation of M-protein occurs. Whilst it may be that turnover of M-like protein is not particularly relevant to this study, due to the previously mentioned associations between M-like protein and HA capsule with regards to virulence and the presence of larger fragments of M-like protein within the mucoid phenotype but not the non-mucoid phenotype, it may not be excessive to suggest a greater link exists between these two virulence factors than has previously been considered. Additionally considering the

variation in sequence observed between the M protein regulator Mga between phenotypes of SE40327 (Chapter 3), this may be an avenue worth further consideration.

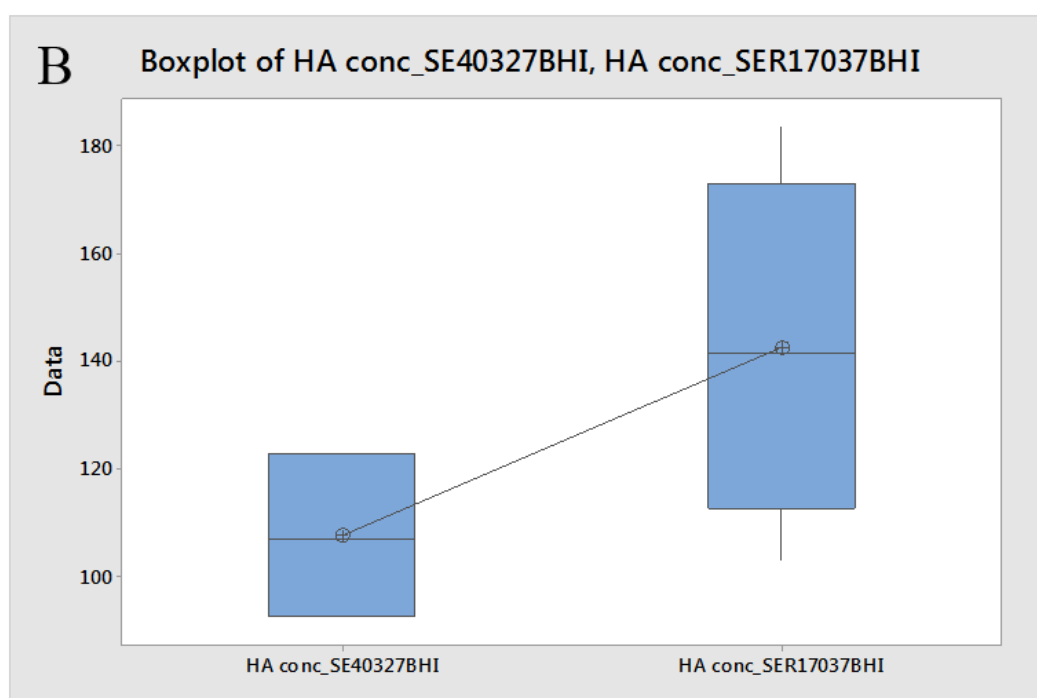
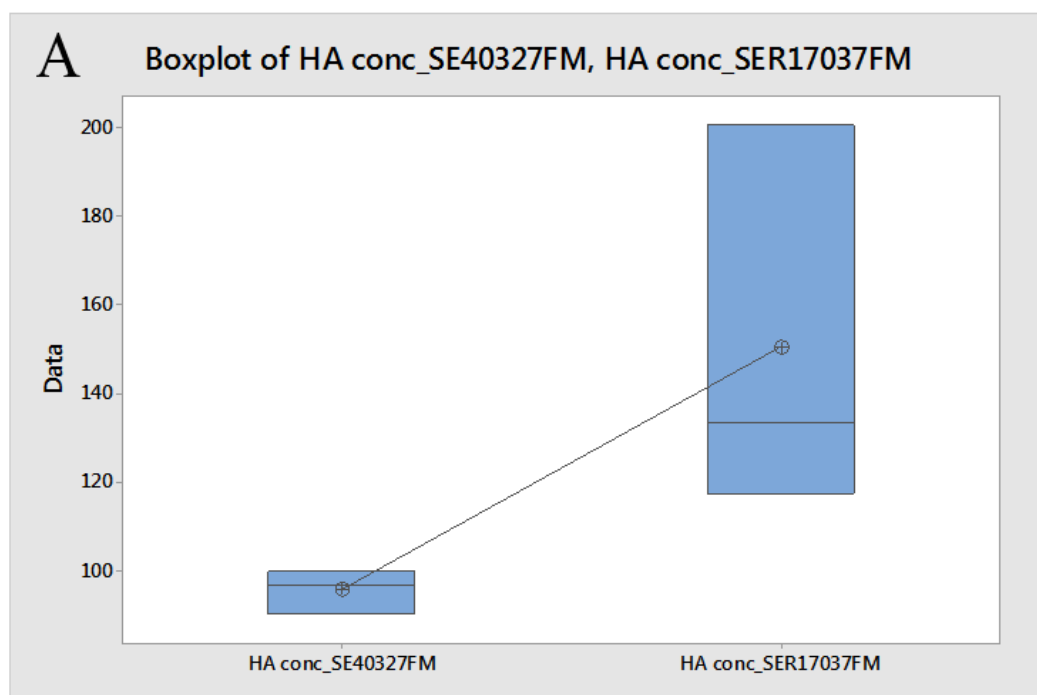
6.4.4. Future work

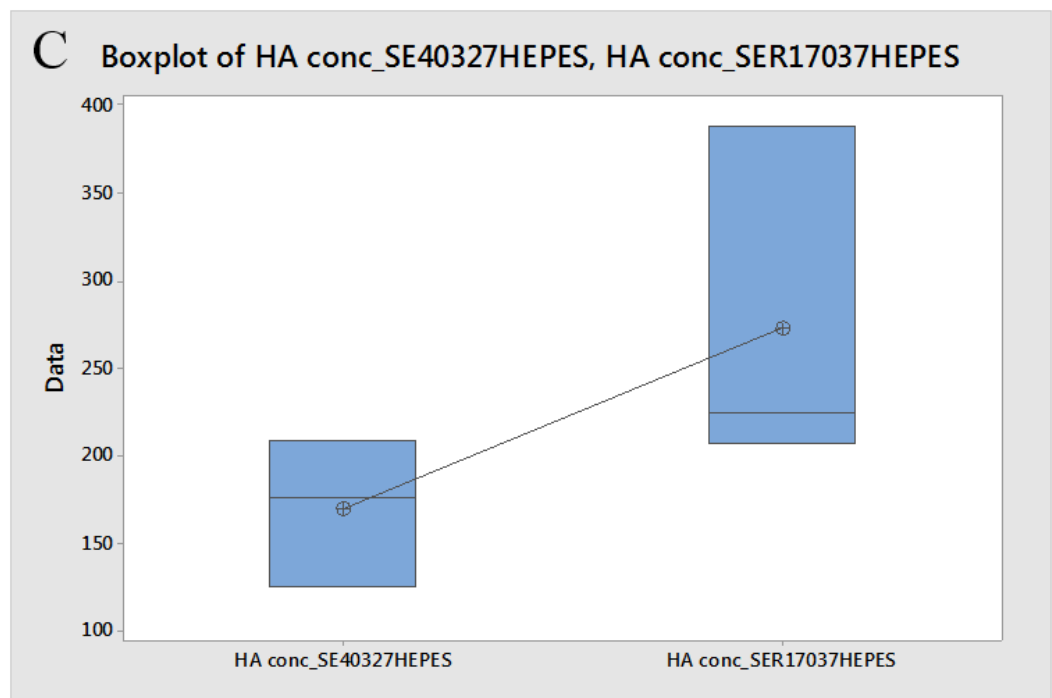
Whilst it is a great first step tool to analysing protein profiles, due to the presence of more than one protein within a band SDS-PAGE often does not reveal all in terms of protein variation between samples. The step up from this technology is Two-Dimensional Difference Gel Electrophoresis (2D-DIGE) which allows for the separation of single proteins and the comparison of more than one sample (Arentz *et al.*, 2015). Whilst this has previously been carried out on HA producing *S. equi* subsp. *zooepidemicus* (Marcellin *et al.*, 2009) the comparison of different phenotypes of the same strain could provide valuable detailed information about protein expression differences relevant to capsule. Higher resolution proteomics using LC-MS/MS as performed with other streptococci (Gao *et al.*, 2019; Qiao *et al.*, 2019) would provide greater coverage in addition to quantitation.

6.5. Overall conclusions

Whilst many studies have been conducted into hyaluronic acid capsule production in Group A and C streptococci, until now the emphasis has either been on the engineering aspect of fermentation parameters or on the genetic machinery but with the narrow scope of the enzymes immediately responsible and their regulation. It would not be arbitrary to suggest that perspective has been lost in terms of recognition of the innate evolutionary reasoning for HA synthesis in these pathogenic bacteria. This study has aimed to return the emphasis to this aspect of HA production research. The results from this study open the scope of research for means to improve yield and characteristics of biomanufactured HA beyond alteration of fermentation parameters and overexpression of the *has* operon.

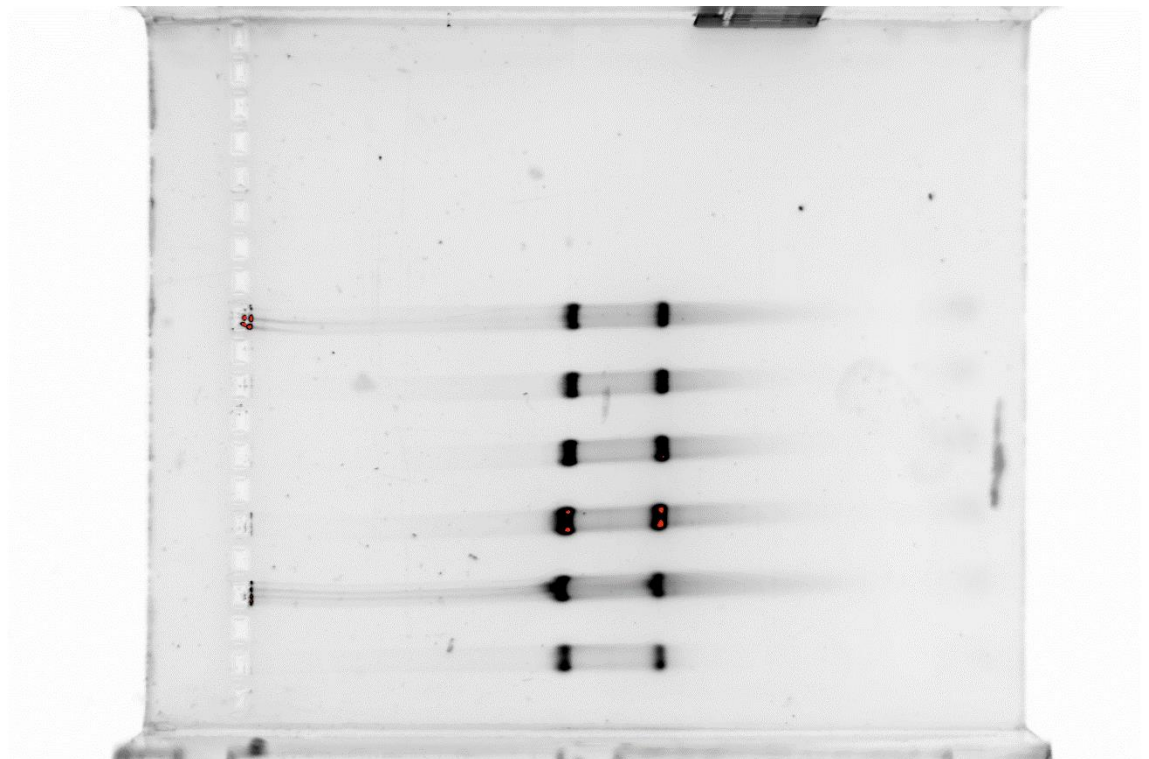
Appendix A. Box plots of HA concentration comparison between SE40327 and SER17037 in Fermentation media (FM), Buffered fermentation media (BFM) and Brain heart infusion media (BHI).





Hyaluronic acid concentration produced by SE40327 compared to SER17037 in fermentation media (A), brain heart infusion media (B) and buffered fermentation media (C).

Appendix B. Agarose gel for RNA sample integrity assessment.



Appendix C. Tables of NCBI database identifications of proteins in bands from SDS-PAGE gel.

SEZ20727-1				
NCBI identification	MOWSE score	% Cover	Strain	No. peptides
ATP-dependent 6-phosphofructokinase	968.70	68.50	<i>S. equi</i>	7
ribonucleotide-diphosphate reductase subunit beta	194.78	7.20	<i>S. equi</i>	2
foldase PrsA	167.99	12.90	<i>S. equi</i>	2
dipeptidase	84.23	6.60	<i>S. equi</i>	2

SER17037-1				
NCBI identification	MOWSE score	% Cover	Strain	No. peptides
peptidase	682.46	20.10	<i>S. equi</i>	8
hypothetical protein	391.49	26.90	<i>S. equi</i>	4
foldase PrsA	126.32	13.2	<i>S. equi</i>	2

SEZ20727-2				
NCBI identification	MOWSE score	% Cover	Strain	No. peptides
superoxide dismutase	599.53	48.30	<i>S. equi</i>	8
hypothetical protein	389.78	37.40	<i>S. equi</i>	5
50S ribosomal protein L4	327.41	41.10	<i>S. equi</i>	2
Uracil phosphoribosyltransferase PyrR	183.99	16.20	<i>S. equi</i>	3
elongation factor Tu	102.89	12.40	<i>S. pneumoniae</i>	2
elongation factor Tu	90.46	8.10	<i>S. pneumoniae</i>	2
hypothetical protein	368.88	17.80	<i>S. equi</i>	3

SER17037-2				
NCBI identification	MOWSE score	% Cover	Strain	No. peptides
hypothetical protein	207.52	11.7	<i>S. equi</i>	2
superoxide dismutase, partial	138.91	17.5	<i>S. zooepidemicus</i>	3
peptidase	785.45	12.9	<i>S. equi</i>	4

SER17037-3				
NCBI identification	MOWSE score	% Cover	Strain	No. peptides
superoxide dismutase	777.26	67.7	<i>S. equi</i>	10
peptidase	756.29	15.4	<i>S. equi</i>	6
ribosome recycling factor	313.31	29.2	<i>S. pyogenes</i>	2
hypothetical protein	190.26	27.9	<i>S. equi</i>	3

SER17037-4				
NCBI identification	MOWSE score	% Cover	Strain	No. peptides
peptidase	728.04	15.20	<i>S. equi</i>	6
30S ribosomal protein S4	295.15	15.30	<i>S. equi</i>	3
hypothetical protein	81.11	8.70	<i>S. equi</i>	2

SER17037-5				
NCBI identification	MOWSE score	% Identity	Strain	No. peptides
superoxide dismutase	1019.76	82.10	<i>S. equi</i>	8
peptidase	631.01	15.10	<i>S. equi</i>	5
50S ribosomal protein L4	428.12	35.70	<i>S. equi</i>	5
transcriptional regulator	358.41	24.60	<i>S. equi</i>	3
ribosome recycling factor	273.77	29.20	<i>S. pyogenes</i>	3
elongation factor Tu	122.15	6.50	<i>S. equi</i>	2
hypoxanthine phosphoribosyltransferase	89.56	11.10	<i>S. equi</i>	2

Appendix D. SER17037 M-like protein amino acid sequence alignment to the sequence identified by Speck *et al.*, (2008) using Multalin.

	1	10	20	30	40	50	60	70	80	90	100																	
SER17037-1	MFLRNKKQKFSIRKL	SAGASVLYATS	SVLGGAVVKA	NISSIT	TNATDRGR	LQAIIDK	GLLNYRL	GEKSE	EQILNDLR	QALSE	ASTDTLATLVSGIDSG																	
CECT5772M																												
Consensus																	
	101	110	120	130	140	150	160	170	180	190	200																	
SER17037-1	RFDYTN	SQFN	YITRN	LGSFRR	DSEDAQ	HDSLN	FASLL	YDEL	VKRL	YSEEL	KGDI	VEKIA	RELE	QRKNA	EVOL	KSEK	ERLE	AQIE	KAGY	IANK	QKELE	KA						
CECT5772M																												
Consensus						
	201	210	220	230	240	250	260	270	280	290	300																	
SER17037-1	RSDQ	KEL	SESI	QKL	TSRF	KKEL	DEAK	VANK	SL	SESV	TKTL	ARSS	KITN	ELK	DKLE	ASEK	DKNR	AIQIT	TEL	TNKL	TDARS	RDKAF	AVYS	KDL	ADKL	AAKT		
CECT5772M																												
Consensus		
	301	310	320	330	340	350	360	370	380	390	400																	
SER17037-1	AEAE	KLMEN	VGS	LDR	LVES	AKRE	AEKL	AEID	QLT	ADK	KAD	AE	LAAN	DTIAS	LQTE	LEKY	KT	LAY	SER	LIES	GKRE	IAEL	QKQ	KDS	KAL	AE	SQAN	
CECT5772M																												
Consensus	aeAa	alnad	sgel	agl	Vas	Akr	dnad	keAa	adkae	ADk	akAd	AdLAA	Aaaa	EId	akad	Lat	ada	Ia	Sge	qe	IAEL	ekki	das	Ka	Lae	aan		
	401	410	420	430	440	450	460	470	480	490	500																	
SER17037-1	VAE	LEKQ	KAES	DAK	VAE	LEKE	VEAK	REY	ADL	KQ	LAK	KEE	ELAV	KKE	EALE	AKIE	EEL	KKAH	AE	SKL	KEH	LEK	DDHAN	ADL	QAE	INRL	KQEL	ADRI
CECT5772M																												
Consensus	saeg	ehqda	enD	aka	aeLek	EeK	EAk	AEI	adIKaa	Lakae	eeLaa	ieeee	Keaaa	AKIda	akKAF	Aeeeee	Khae	aLa	afae	AFa	Alka	ika	ekka	aaa	aard			
	501	510	520	530	540	550	558																					
SER17037-1	KSL	SQSG	RAS	QTNS	GTTT	AKAG	QLP	STGES	ANP	FFT	IA	RLTY	IAG	AGM	AVYS	PKR	KEN											
CECT5772M																												
Consensus	fsl	kkidr	An	qtd	ggk	Tlae	agld	piage	ee	Pe	akia	akpe	ia	gag	kaep	kP											

Antiphagocytic M-like protein identified in this study (SER17037-1) compared to M-like protein identified previously (CECT5772M), NCBI accession number (EU069409.1).

References

- Abbott, Y., Acke, E., Khan, S., Muldoon, E.G., Markey, B.K., Pinilla, M., Leonard, F.C., Steward, K. and Waller, A. (2010) 'Zoonotic transmission of *Streptococcus equi* subsp. *zooepidemicus* from a dog to a handler', *J Med Microbiol*, 59(Pt 1), pp. 120-3.
- Ali, L., Blum, H.E. and Sakiotanc, T. (2019) 'Detection and characterization of bacterial polysaccharides in drug-resistant enterococci', *Glycoconj J*, 36(5), pp. 429-438.
- Alkan, C., Sajjadian, S. and Eichler, E.E. (2011) 'Limitations of next-generation genome sequence assembly', *Nat Methods*, 8(1), pp. 61-5.
- Ambrosio, L., Borzacchiello, A., Netti, P.A. and Nicolais, L. (1999) 'Rheological Study on Hyaluronic Acid and Its Derivative Solutions', *Journal of Macromolecular Science, Part A*, 36(7-8), pp. 991-1000.
- Anzai, T., Timoney, J.F., Kuwamoto, Y., Fujita, Y., Wada, R. and Inoue, T. (1999) 'In vivo pathogenicity and resistance to phagocytosis of *Streptococcus equi* strains with different levels of capsule expression', *Veterinary Microbiology*, 67, pp. 277-286.
- Arentz, G., Weiland, F., Oehler, M.K. and Hoffmann, P. (2015) 'State of the art of 2D DIGE', *PROTEOMICS - Clinical Applications*, 9(3-4), pp. 277-288.
- Armstrong, D.C. and Johns, M.R. (1997) 'Culture Conditions Affect the Molecular Weight Properties of Hyaluronic Acid Produced by *Streptococcus zooepidemicus*', *Applied and Environmental Microbiology*, 63, pp. 2759-2764.
- Armstrong, D.C., Cooney, M.J. and Johns, M.R. (1997) 'Growth and amino acid requirements of hyaluronic-acid-producing *Streptococcus zooepidemicus*', *Appl Microbiol Biotechnol*, 47, pp. 309-312.
- Arora, J., Sauer, S.J., Tarpley, M., Vermeulen, P., Rypens, C., Van Laere, S., Williams, K.P., Devi, G.R. and Dewhirst, M.W. (2017) 'Inflammatory breast cancer tumor emboli express high levels of anti-apoptotic proteins: use of a quantitative high content and high-throughput 3D IBC spheroid assay to identify targeting strategies', *Oncotarget*, 8(16), pp. 25848-25863.
- Asari, A., Kanemitsu, T. and Kurihara, H. (2010) 'Oral administration of high molecular weight hyaluronan (900 kDa) controls immune system via Toll-like receptor 4 in the intestinal epithelium', *J Biol Chem*, 285(32), pp. 24751-8.
- Aya, K.L. and Stern, R. (2014) 'Hyaluronan in wound healing: Rediscovering a major player', *Wound Repair and Regeneration*, 22, pp. 579-593.
- Aya, K.L. and Stern, R. (2013) 'Hyaluronan in wound healing: Rediscovering a major player', *Wound Repair and Regeneration*, 22(579-593).

Badri, A., Williams, A., Linhardt, R.J. and Koffas, M.A. (2017) 'The road to animal-free glycosaminoglycan production: current efforts and bottlenecks', *Curr Opin Biotechnol*, 53, pp. 85-92.

Balazs, E.A. (2009) 'Therapeutic use of hyaluronan', *Structural Chemistry*, 20(2), pp. 341-349.

Bannoehr, J., Ben Zakour, N.L., Reglinski, M., Inglis, N.F., Prabhakaran, S., Fossum, E., Smith, D.G., Wilson, G.J., Cartwright, R.A., Haas, J., Hook, M., van den Broek, A.H., Thoday, K.L. and Fitzgerald, J.R. (2011) 'Genomic and surface proteomic analysis of the canine pathogen *Staphylococcus pseudintermedius* reveals proteins that mediate adherence to the extracellular matrix', *Infect Immun*, 79(8), pp. 3074-86.

Barreteau, H., Kovac, A., Boniface, A., Sova, M., Gobec, S. and Blanot, D. (2008) 'Cytoplasmic steps of peptidoglycan biosynthesis', *FEMS Microbiol Rev*, 32(2), pp. 168-207.

Batycka, M., Inglis, N.F., Cook, K., Adam, A., Fraser-Pitt, D., Smith, D.G., Main, L., Lubben, A. and Kessler, B.M. (2006) 'Ultra-fast tandem mass spectrometry scanning combined with monolithic column liquid chromatography increases throughput in proteomic analysis', *Rapid Commun Mass Spectrom*, 20(14), pp. 2074-80.

Benedini, L.J. and Santana, M.H. (2013) 'Effects of soy peptone on the inoculum preparation of *Streptococcus zooepidemicus* for production of hyaluronic acid', *Bioresour Technol*, 130, pp. 798-800.

Bhatia, S. (2016) 'Mammalian Polysaccharides and Its Nanomaterials', in *Systems for Drug Delivery: Safety, Animal, and Microbial Polysaccharides*. Cham: Springer International Publishing, pp. 1-27.

Bi, Y., Hubbard, C., Purushotham, P. and Zimmer, J. (2015) 'Insights into the structure and function of membrane-integrated processive glycosyltransferases', *Curr Opin Struct Biol*, 34, pp. 78-86.

Bioline (2015) 'One-step Vs. Two-step real-time RT PCR', *BIOLINE Blog - Real-Time PCR*. Available at: <https://www.bioline.com/us/one-step-vs-two-step-real-time-pcr.html>.

Blank, L.M., McLaughlin, R.L. and Nielsen, L.K. (2005) 'Stable production of hyaluronic acid in *Streptococcus zooepidemicus* chemostats operated at high dilution rate', *Biotechnol Bioeng*, 90(6), pp. 685-93.

Blank, L.M., Hugenholtz, P. and Nielsen, L.K. (2008) 'Evolution of the hyaluronic acid synthesis (has) operon in *Streptococcus zooepidemicus* and other pathogenic streptococci', *J Mol Evol*, 67(1), pp. 13-22.

- Boeriu, C.G., Springer, J., Kooy, F.K., van den Broek, L.A.M. and Eggink, G. (2013) 'Production Methods for Hyaluronan', *International Journal of Carbohydrate Chemistry*, 2013, pp. 1-14.
- Boyle, A.G., Rankin, S.C., Duffee, L., Boston, R.C. and Wheeler-Aceto, H. (2016) 'Streptococcus equi Detection Polymerase Chain Reaction Assay for Equine Nasopharyngeal and Guttural Pouch Wash Samples', *J Vet Intern Med*, 30(1), pp. 276-81.
- Brem, H. and Tomic-Canic, M. (2007) 'Cellular and molecular basis of wound healing in diabetes', *J Clin Invest*, 117(5), pp. 1219-22.
- Brouwer, S., Barnett, T.C., Rivera-Hernandez, T., Rohde, M. and Walker, M.J. (2016) 'Streptococcus pyogenes adhesion and colonization', *FEBS Lett*, 590(21), pp. 3739-3757.
- Bustin, S. and Huggett, J. (2017) 'qPCR primer design revisited', *Biomol Detect Quantif*, 14, pp. 19-28.
- Bustin, S. (2017) 'The continuing problem of poor transparency of reporting and use of inappropriate methods for RT-qPCR', *Biomol Detect Quantif*, 12, pp. 7-9.
- Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., Vandesompele, J. and Wittwer, C.T. (2009) 'The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments', *Clin Chem*, 55(4), pp. 611-22.
- Chakraborty, I., Sen, I.K., Mondal, S., Rout, D., Bhanja, S.K., Maity, G.N. and Maity, P. (2019) 'Bioactive polysaccharides from natural sources: A review on the antitumor and immunomodulating activities', *Biocatalysis and Agricultural Biotechnology*, 22.
- Chanmee, T., Ontong, P. and Itano, N. (2016) 'Hyaluronan: A modulator of the tumor microenvironment', *Cancer Lett*, 375(1), pp. 20-30.
- Chanter, N., Ward, C.L., Talbot, N.C., Flanagan, J.A., Binns, M., Houghton, S.B., Smith, K.C. and Mumford, J.A. (1999) 'Recombinant hyaluronate associated protein as a protective immunogen against Streptococcus equi and Streptococcus zooepidemicus challenge in mice', *Microbial Pathogenesis*, 27, pp. 133-143.
- Chassy, B.M. and Giuffrida, A. (1980) 'Method for the Lysis of Gram-Positive, Asporogenous Bacteria with Lysozyme', *Appl Environ Microbiol*, 39(1), pp. 153-158.
- Chauhan, A.S., Badle, S.S., Ramachandran, K.B. and Jayaraman, G. (2014) 'The P170 expression system enhances hyaluronan molecular weight and production in

- metabolically-engineered *Lactococcus lactis*', *Biochemical Engineering Journal*, 90, pp. 73-78.
- Chaussee, M.A., McDowell, E.J., Rieck, L.D., Callegari, E.A. and Chaussee, M.S. (2006) 'Proteomic analysis of a penicillin-tolerant rgg mutant strain of *Streptococcus pyogenes*', *J Antimicrob Chemother*, 58(4), pp. 752-9.
- Chen, S.-J., Chen, J.-L., Huang, W.-C. and Chen, H.-L. (2009a) 'Fermentation process development for hyaluronic acid production by *Streptococcus zooepidemicus* ATCC 39920', *Korean Journal of Chemical Engineering*, 2, pp. 428-432.
- Chen, W.Y., Marcellin, E., Steen, J.A. and Nielsen, L.K. (2014) 'The role of hyaluronic acid precursor concentrations in molecular weight control in *Streptococcus zooepidemicus*', *Mol Biotechnol*, 56(2), pp. 147-56.
- Chen, W.Y., Marcellin, E., Hung, J. and Nielsen, L.K. (2009b) 'Hyaluronan molecular weight is controlled by UDP-N-acetylglucosamine concentration in *Streptococcus zooepidemicus*', *J Biol Chem*, 284(27), pp. 18007-14.
- Chen, Y.-H., Li, J., Liu, L., Liu, H.-Z. and Wang, Q. (2012) 'Optimization of Flask Culture Medium and Conditions for Hyaluronic Acid Production by a *Streptococcus Equisimilis* Mutant NC2168', *Brazilian Journal of Microbiology*, pp. 1553-1561.
- Chiang-Ni, C., Nian, S.Y., Wu, J.J. and Chen, C.J. (2016) 'Oxygen-dependent phenotypic variation in group A streptococcus', *J Microbiol Immunol Infect*, 49(6), pp. 837-842.
- Choi, S., Choi, W., Kim, S., Lee, S.-Y., Noh, I. and Kim, C.-W. (2014) 'Purification and biocompatibility of fermented hyaluronic acid for its applications to biomaterials', *Biomaterials Research*, 18(6).
- Chong, B.F., Blank, L.M., McLaughlin, R. and Nielsen, L.K. (2005) 'Microbial hyaluronic acid production', *Appl Microbiol Biotechnol*, 66(4), pp. 341-51.
- Cleary, P.P. and Larkin, A. (1979) 'Hyaluronic Acid Capsule: Strategy for Oxygen Resistance in Group A Streptococci', *Journal of Bacteriology*, 140(3), pp. 1090-1097.
- Cleary, P.P., McLandsborough, L., Ikeda, L., Cue, D., Krawczak, J. and Lam, H. (1998) 'High-frequency intracellular infection and erythrogenic toxin A expression undergo phase variation in M1 group A streptococci', *Molecular Microbiology*, 28(1), pp. 157-167.
- Cole, J.N., Pence, M.A., von Kockritz-Blickwede, M., Hollands, A., Gallo, R.L., Walker, M.J. and Nizet, V. (2010) 'M protein and hyaluronic acid capsule are essential

for in vivo selection of covRS mutations characteristic of invasive serotype M1T1 group A *Streptococcus*', *MBio*, 1(4).

Cowman, M.K., Schmidt, T.A., Raghavan, P. and Stecco, A. (2015) 'Viscoelastic Properties of Hyaluronan in Physiological Conditions', *F1000Res*, 4, p. 622.

Cress, B.F., Englaender, J.A., He, W., Kasper, D., Linhardt, R.J. and Koffas, M.A. (2014) 'Masquerading microbial pathogens: capsular polysaccharides mimic host-tissue molecules', *FEMS Microbiol Rev*, 38(4), pp. 660-97.

Crowley, P.J. and Brady, L.J. (2016) 'Evaluation of the effects of *Streptococcus mutans* chaperones and protein secretion machinery components on cell surface protein biogenesis, competence, and mutacin production', *Mol Oral Microbiol*, 31(1), pp. 59-77.

Cyphert, J.M., Trempus, C.S. and Garantziotis, S. (2015) 'Size Matters: Molecular Weight Specificity of Hyaluronan Effects in Cell Biology', *Int J Cell Biol*, 2015, p. 563818.

Dale, J.B., Washburn, R.G., Marques, M.B. and Wessels, M.R. (1996a) 'Hyaluronate Capsule and Surface M Protein in Resistance to Opsonization of Group A *Streptococci*', *Infection and Immunity*, 64(5), pp. 1495-1501.

Dale, J.B., Washburn, R.G., Marques, M.B. and Wessels, M.R. (1996b) 'Hyaluronate Capsule and Surface M Protein in Resistance to Opsonization of Group A *Streptococci*', *Infection and Immunity*, 64, pp. 1495-1501.

Denn, M.M. (2004) 'Fifty years of non-Newtonian fluid dynamics', *AIChE Journal*, 50(10), pp. 2335-2345.

Denton, J.F., Lugo-Martinez, J., Tucker, A.E., Schrider, D.R., Warren, W.C. and Hahn, M.W. (2014) 'Extensive error in the number of genes inferred from draft genome assemblies', *PLoS Comput Biol*, 10(12), p. e1003998.

Dierksen, K.P., Ragland, N.I. and Tagg, J.R. (2000) 'A New Alkaline pH-Adjusted Medium Enhances Detection of β -Hemolytic *Streptococci* by Minimizing Bacterial Interference Due to *Streptococcus salivarius*', *Journal of Clinical Microbiology*, 38, pp. 643-650.

Dinkla, K., Sastalla, I., Godehardt, A.W., Janze, N., Chhatwal, G.S., Rohde, M. and Medina, E. (2007) 'Upregulation of capsule enables *Streptococcus pyogenes* to evade immune recognition by antigen-specific antibodies directed to the G-related alpha2-macroglobulin-binding protein GRAB located on the bacterial surface', *Microbes Infect*, 9(8), pp. 922-31.

- Don, M.M. and Shoparwe, N.F. (2010) 'Kinetics of hyaluronic acid production by *Streptococcus zooepidemicus* considering the effect of glucose', *Biochemical Engineering Journal*, 49(1), pp. 95-103.
- Dosio, F., Arpicco, S., Stella, B. and Fattal, E. (2016) 'Hyaluronic acid for anticancer drug and nucleic acid delivery', *Adv Drug Deliv Rev*, 97, pp. 204-36.
- Du, W., Brown, J.R., Sylvester, D.R., Huang, J., Chalker, A.F., So, C.Y., Holmes, D.J., Payne, D.J. and Wallis, N.G. (2000) 'Two Active Forms of UDP-N-Acetylglycosamine Enolpyruvyl Transferase in Gram-Positive Bacteria', *Journal of Bacteriology*, 182, pp. 4146-4152.
- Duan, Xu-Jie, Yang, L., Zhang, X. and Tan, W.-S. (2008) 'Effect of Oxygen and Shear Stress on Molecular Weight of Hyaluronic Acid Produced by *Streptococcus zooepidemicus*', *Journal of Microbiology and Biotechnology*, 18(4), pp. 718-724.
- Emerson, D., Agulto, L., Liu, H. and Liu, L. (2008) 'Identifying and Characterizing Bacteria in an Era of Genomics and Proteomics', *Bioscience*, 58, pp. 925-936.
- Engel, H., Gutierrez-Fernandez, J., Fluckiger, C., Martinez-Ripoll, M., Muhlemann, K., Hermoso, J.A., Hilty, M. and Hathaway, L.J. (2013) 'Heteroresistance to Fosfomycin Is Predominant in *Streptococcus pneumoniae* and Depends on the *murA1* Gene', *Antimicrobial Agents and Chemotherapy*, 57, pp. 2801-2808.
- EnvironmentAgency (2018) *Significant fine for water company following hazardous chemical leak*. Available at: <https://www.gov.uk/government/news/significant-fine-for-water-company-following-hazardous-chemical-leak> (Accessed: 22 May 2019).
- Eschenburg, S., Priestman, M.A., Abdul-Latif, F.A., Delachaume, C., Fassy, F. and Schonbrunn, E. (2005) 'A novel inhibitor that suspends the induced fit mechanism of UDP-N-acetylglycosamine enolpyruvyl transferase (MurA)', *J Biol Chem*, 280(14), pp. 14070-5.
- Falagas, M.E., Vouloumanou, E.K., Samonis, G. and Vardakas, K.Z. (2016) 'Fosfomycin', *Clin Microbiol Rev*, 29(2), pp. 321-47.
- Falaleeva, M., Zurek, O.W., Watkins, R.L., Reed, R.W., Ali, H., Sumby, P., Voyich, J.M. and Korotkova, N. (2014) 'Transcription of the *Streptococcus pyogenes* Hyaluronic Acid Capsule Biosynthesis Operon Is Regulated by Previously Unknown Upstream Elements', *Infection and Immunity*, 82(12), pp. 5293 - 5307.
- Fallacara, A., Baldini, E., Manfredini, S. and Vertuani, S. (2018) 'Hyaluronic Acid in the Third Millennium', *Polymers (Basel)*, 10(7).

Fernandez, E., Blume, V., Garrido, P., Collins, M.D., Mateos, A., Dominguez, L. and Fernandez-Garayzabal, J.F. (2004) 'Streptococcus equi subsp. ruminatorum subsp. nov., isolated from mastitis in small ruminants', *Int J Syst Evol Microbiol*, 54(Pt 6), pp. 2291-6.

Ferrandiz, M.J., Cercenado, M.I., Domenech, M., Tirado-Velez, J.M., Escolano-Martinez, M.S., Yuste, J., Garcia, E., de la Campa, A.G. and Martin-Galiano, A.J. (2019) 'An Uncharacterized Member of the Gls24 Protein Superfamily Is a Putative Sensor of Essential Amino Acid Availability in *Streptococcus pneumoniae*', *Microb Ecol*, 77(2), pp. 471-487.

Field, T.R., Ward, P.N., Pedersen, L.H. and Leigh, J.A. (2003) 'The hyaluronic acid capsule of *Streptococcus uberis* is not required for the development of infection and clinical mastitis', *Infect Immun*, 71(1), pp. 132-9.

Fischetti, V.A. (2016) 'M Protein and Other Surface Proteins on Streptococci', in Ferretti, J., Stevens, D. and Fischetti, V. (eds.) *Streptococcus pyogenes: Basic Biology to Clinical Manifestations*. The University of Oklahoma Health Sciences Center.

Fischetti, V.A. and Dale, J.B. (2016) 'One More Disguise in the Stealth Behavior of *Streptococcus pyogenes*', *MBio*, 7(3).

Flores, S.C., Lu, L.J., Yang, J., Carriero, N. and Gerstein, M.B. (2007) 'Hinge Atlas: relating protein sequence to sites of structural flexibility', *BMC Bioinformatics*, 8, p. 167.

Fouda, M.M., Abdel-Mohsen, A.M., Ebaid, H., Hassan, I., Al-Tamimi, J., Abdel-Rahman, R.M., Metwalli, A., Alhazza, I., Rady, A., El-Faham, A. and Jancar, J. (2016) 'Wound healing of different molecular weight of hyaluronan; in-vivo study', *Int J Biol Macromol*, 89, pp. 582-91.

Gallo, N., Nasser, H., Salvatore, L., Natali, M.L., Campa, L., Mahmoud, M., Capobianco, L., Sannino, A. and Madaghiele, M. (2019) 'Hyaluronic acid for advanced therapies: Promises and challenges', *European Polymer Journal*, 117, pp. 134-147.

Gao, N.J., Al-Bassam, M.M., Poudel, S., Wozniak, J.M., Gonzalez, D.J., Olson, J., Zengler, K., Nizet, V. and Valderrama, J.A. (2019) 'Functional and Proteomic Analysis of *Streptococcus pyogenes* Virulence Upon Loss of Its Native Cas9 Nuclease', *Front Microbiol*, 10, p. 1967.

Gao, X.Y., Zhi, X.Y., Li, H.W., Klenk, H.P. and Li, W.J. (2014) 'Comparative genomics of the bacterial genus *Streptococcus* illuminates evolutionary implications of species groups', *PLoS One*, 9(6), p. e101229.

Ghosh, P. (2018) 'Variation, Indispensability, and Masking in the M protein', *Trends Microbiol*, 26(2), pp. 132-144.

Giard, J.-C., Verneuil, N., Auffray, Y. and Hartke, A. (2002) 'Characterization of genes homologous to the general stress-inducible gene *gls24* in *Enterococcus faecalis* and *Lactococcus lactis* ', *FEMS Microbiol Lett*, 206, pp. 235-239.

Girish, K.S. and Kemparaju, K. (2007) 'The magic glue hyaluronan and its eraser hyaluronidase: a biological overview', *Life Sci*, 80(21), pp. 1921-43.

Gottesman, S. (2019) 'Trouble is coming: Signaling pathways that regulate general stress responses in bacteria', *J Biol Chem*, 294(31), pp. 11685-11700.

Guo, L., Wu, T., Hu, W., He, X., Sharma, S., Webster, P., Gimzewski, J.K., Zhou, X., Lux, R. and Shi, W. (2012) 'Phenotypic characterization of the foldase homologue PrsA in *Streptococcus mutans*', *Mol Oral Microbiol*, 28(2), pp. 154-65.

Hamilton, S., McLean, K., Monaghan, S.J., McNair, C., Inglis, N.F., McDonald, H., Adams, S., Richards, R., Roy, W., Smith, P., Bron, J., Nisbet, A.J. and Knox, D. (2018) 'Characterisation of proteins in excretory/secretory products collected from salmon lice, *Lepeophtheirus salmonis*', *Parasit Vectors*, 11(1), p. 294.

Han, H., Yang, Y., Olesen, S.H., Becker, A., Betzi, S. and Schonbrunn, E. (2010) 'The fungal product terreic acid is a covalent inhibitor of the bacterial cell wall biosynthetic enzyme UDP-N-acetylglucosamine 1-carboxyvinyltransferase (MurA)', *Biochemistry*, 49(19), pp. 4276-82.

Harrington, D.J., Sutcliffe, I.C. and Chanter, N. (2002) 'The molecular basis of *Streptococcus equi* infection and disease', *Microbes and Infection*, 4, pp. 501-510.

Hartke, A., Giard, J.-C., Laplace, J.-M. and Auffray, Y. (1998) 'Survival of *Enterococcus faecalis* in an Oligotrophic Microcosm: Changes in Morphology, Development of General Stress Resistance, and Analysis of Protein Synthesis', *Applied and Environmental Microbiology*, 64(11), pp. 4238-4245.

Heldermon, C., DeAngelis, P.L. and Weigel, P.H. (2001) 'Topological organization of the hyaluronan synthase from *Streptococcus pyogenes*', *J Biol Chem*, 276(3), pp. 2037-46.

Heldermon, C.D., Tlapak-Simmons, V.L., Baggenstoss, B.A. and Weigel, P.H. (2001) 'Site-directed mutation of conserved cysteine residues does not inactivate the *Streptococcus pyogenes* hyaluronan synthase', *Glycobiology*, 11, pp. 1017-1024.

Hemshekhhar, M., Thushara, R.M., Chandranayaka, S., Sherman, L.S., Kemparaju, K. and Girish, K.S. (2016) 'Emerging roles of hyaluronic acid bioscaffolds in tissue engineering and regenerative medicine', *Int J Biol Macromol*, 86, pp. 917-28.

Henningham, A., Dohrmann, S., Nizet, V. and Cole, J.N. (2015) 'Mechanisms of group A *Streptococcus* resistance to reactive oxygen species', *FEMS Microbiol Rev*, 39(4), pp. 488-508.

Holden, M.T., Heather, Z., Paillot, R., Steward, K.F., Webb, K., Ainslie, F., Jourdan, T., Bason, N.C., Holroyd, N.E., Mungall, K., Quail, M.A., Sanders, M., Simmonds, M., Willey, D., Brooks, K., Aanensen, D.M., Spratt, B.G., Jolley, K.A., Maiden, M.C., Kehoe, M., Chanter, N., Bentley, S.D., Robinson, C., Maskell, D.J., Parkhill, J. and Waller, A.S. (2009) 'Genomic evidence for the evolution of *Streptococcus equi*: host restriction, increased virulence, and genetic exchange with human pathogens', *PLoS Pathog*, 5(3), p. e1000346.

Honer, O.P., Wachter, B., Speck, S., Wibbelt, G., Ludwig, A., Fyumagwa, R.D., Wohlsein, P., Lieckfeldt, D., Hofer, H. and East, M.L. (2006) 'Severe *Streptococcus* infection in spotted hyenas in the Ngorongoro Crater, Tanzania', *Vet Microbiol*, 115(1-3), pp. 223-8.

Hoskisson, P.A. and Hobbs, G. (2005) 'Continuous culture--making a comeback?', *Microbiology*, 151(Pt 10), pp. 3153-9.

Hrast, M., Susic, I., Sink, R. and Gobec, S. (2014) 'Inhibitors of the peptidoglycan biosynthesis enzymes MurA-F', *Bioorg Chem*, 55, pp. 2-15.

Huang, W.-C., Chen, S.-J. and Chen, T.-L. (2008) 'Production of hyaluronic acid by repeated batch fermentation', *Biochemical Engineering Journal*, 40(3), pp. 460-464.

Huang, W.-C., Chen, S.-J. and Chen, T.-L. (2006) 'The role of dissolved oxygen and function of agitation in hyaluronic acid fermentation', *Biochemical Engineering Journal*, 32(3), pp. 239-243.

Hynes, W. (2000) 'Hyaluronidases of Gram-positive bacteria', *FEMS Microbiology Letters*, 183(2), pp. 201-207.

Hynes, W.L. and Walton, S.L. (2000) 'Hyaluronidases of Gram-positive bacteria', *FEMS Microbiology Letters*, 183, pp. 201-207.

Hynes, W.L., Hancock, L. and Ferretti, J. (1995) 'Analysis of a Second Bacteriophage Hyaluronidase Gene from *Streptococcus pyogenes*: Evidence for a Third Hyaluronidase Involved in Extracellular Enzymatic Activity', *Infection and Immunity*, 63(8), pp. 3015-3020.

Hynes, W.L., Dixon, A.R., Walton, S.L. and Aridgides, L.J. (2000) 'The extracellular hyaluronidase gene (*hylA*) of *Streptococcus pyogenes*', *FEMS Microbiol Lett*, 184, pp. 109-112.

- Iannitti, T., Lodi, D. and Palmieri, B. (2011) 'Intra-Articular Injections for the Treatment of Osteoarthritis', *Drugs R D*, 1, pp. 13-27.
- Im, G.I. (2019) 'Perspective on Intra-articular Injection Cell Therapy for Osteoarthritis Treatment', *Tissue Eng Regen Med*, 16(4), pp. 357-363.
- Jakob, R.P., Koch, J.R., Burmann, B.M., Schmidpeter, P.A., Hunkeler, M., Hiller, S., Schmid, F.X. and Maier, T. (2015) 'Dimeric Structure of the Bacterial Extracellular Foldase PrsA', *J Biol Chem*, 290(6), pp. 3278-92.
- Jalali, M., Zaborowska, J. and Jalali, M. (2017) 'The Polymerase Chain Reaction', in *Basic Science Methods for Clinical Researchers*. pp. 1-18.
- Jiang, D., Liang, J. and Noble, P.W. (2011) 'Hyaluronan as an immune regulator in human diseases', *Physiol Rev*, 91(1), pp. 221-64.
- Jores, J., Schieck, E., Liljander, A., Sacchini, F., Posthaus, H., Lartigue, C., Blanchard, A., Labroussaa, F. and Vashee, S. (2019) 'In vivo role of capsular polysaccharide in *Mycoplasma mycoides*', *J Infect Dis*, 219(10), pp. 1559-1563.
- Jukic, M., Gobec, S. and Sova, M. (2019) 'Reaching toward underexplored targets in antibacterial drug design', *Drug Dev Res*, 80(1), pp. 6-10.
- Takehi, K., Kinoshita, M. and Yasueda, S.-i. (2003) 'Hyaluronic acid: separation and biological implications', *Journal of Chromatography B*, 797(1-2), pp. 347-355.
- Kang, S.O., Wright, J.O., Tesorero, R.A., Lee, H., Beall, B. and Cho, K.H. (2012) 'Thermoregulation of capsule production by *Streptococcus pyogenes*', *PLoS One*, 7(5), p. e37367.
- Kaur, M. and Jayaraman, G. (2016) 'Hyaluronan production and molecular weight is enhanced in pathway-engineered strains of lactate dehydrogenase-deficient *Lactococcus lactis*', *Metabolic Engineering Communications*, 3, pp. 15-23.
- Kazmierczak, M.J., Wiedmann, M. and Boor, K.J. (2005) 'Alternative sigma factors and their roles in bacterial virulence', *Microbiol Mol Biol Rev*, 69(4), pp. 527-43.
- Kedar, G.C., Brown-Driver, V., Reyes, D.R., Hilgers, M.T., Stidham, M.A., Shaw, K.J., Finn, J. and Haselbeck, R.J. (2008) 'Comparison of the essential cellular functions of the two *murA* genes of *Bacillus anthracis*', *Antimicrob Agents Chemother*, 52(6), pp. 2009-13.
- Kelly, C., Bugg, M., Robinson, C., Mitchell, Z., Davis-Poynter, N., Newton, J.R., Jolley, K.A., Maiden, M.C. and Waller, A.S. (2006) 'Sequence variation of the *SeM*

gene of *Streptococcus equi* allows discrimination of the source of strangles outbreaks', *J Clin Microbiol*, 44(2), pp. 480-6.

Kim, H., Jeong, H., Han, S., Beack, S., Hwang, B.W., Shin, M., Oh, S.S. and Hahn, S.K. (2017) 'Hyaluronate and its derivatives for customized biomedical applications', *Biomaterials*, 123, pp. 155-171.

Kim, J.-H., Yoo, S.-J., Oh, D.-K., Kweon, Y.-G., Park, D.-W., Lee, C.-H. and Gil, G.-H. (1996) 'Selection of a *Streptococcus equi* mutant and optimization of culture conditions for the productions of high molecular weight hyaluronic acid', *Enzyme and Microbial Technology*, 19, pp. 440-445.

Kim, S.-J., Park, S.-Y. and Kim, C.-W. (2006) 'A Novel Approach to the Production of Hyaluronic Acid by *Streptococcus zooepidemicus*', *Journal of Microbiology and Biotechnology*, 16(12), pp. 1849-1855.

Klein, C.D. and Bachelier, A. (2006) 'Molecular modeling and bioinformatical analysis of the antibacterial target enzyme MurA from a drug design perspective', *J Comput Aided Mol Des*, 20(10-11), pp. 621-8.

Kleinberg, T.T., Tzekov, R.T., Stein, L., Ravi, N. and Kaushal, S. (2011) 'Vitreous substitutes: a comprehensive review', *Surv Ophthalmol*, 56(4), pp. 300-23.

Klos, M. and Wojkowska-Mach, J. (2017) 'Pathogenicity of Virulent Species of Group C *Streptococci* in Human', *Can J Infect Dis Med Microbiol*, 2017, p. 9509604.

Knudson, W., Ishizuka, S., Terabe, K., Askew, E.B. and Knudson, C.B. (2019) 'The pericellular hyaluronan of articular chondrocytes', *Matrix Biol*, 78-79, pp. 32-46.

Kogan, G., Soltes, L., Stern, R. and Gemeiner, P. (2007) 'Hyaluronic acid: a natural biopolymer with a broad range of biomedical and industrial applications', *Biotechnol Lett*, 29(1), pp. 17-25.

Krzyściak, W., Pluskwa, K.K., Jurczak, A. and Kościelniak, D. (2013) 'The pathogenicity of the *Streptococcus* genus', *Eur J Clin Microbiol Infect Dis*, 32(11), pp. 1361-76.

Lamagni, T.L., Darenberg, J., Luca-Harari, B., Siljander, T., Efstratiou, A., Henriques-Normark, B., Vuopio-Varkila, J., Bouvet, A., Creti, R., Ekelund, K., Koliou, M., Reinert, R.R., Stathi, A., Strakova, L., Ungureanu, V., Schalen, C., Strep, E.S.G. and Jasir, A. (2008) 'Epidemiology of severe *Streptococcus pyogenes* disease in Europe', *J Clin Microbiol*, 46(7), pp. 2359-67.

Lee, J.Y. and Spicer, A.P. (2000) 'Hyaluronan: a multifunctional, megaDalton stealth molecule', *Current Opinion in Cell Biology*, 12, pp. 581-586.

- Lee, M.-Y., Yang, J.-A., Jung, H.S., Beack, S., Choi, J.E., Hur, W., Koo, H., Kim, K., Yoon, S.K. and Hahn, S.K. (2012) 'Hyaluronic Acid-Gold Nanoparticle/Interferon Alpha Complex for Targeted Treatment of Hepatitis C Virus Infection', *ACS Nano*, 6, pp. 9522-9531.
- Leonard, B.A.B., Woischnik, M. and Podbielski, A. (1998) 'Production of Stabilized Virulence Factor-negative Variants by Group A Streptococci during Stationary Phase', *Infection and Immunity*, 66(8), pp. 3841-3847.
- Li, W., Yi, X., Liu, X., Zhang, Z., Fu, Y. and Gong, T. (2016) 'Hyaluronic acid ion-pairing nanoparticles for targeted tumor therapy', *J Control Release*, 225, pp. 170-82.
- Libardoni, F., Vielmo, A., Farias, L., Matter, L.B., Potter, L., Spilki, F.R. and de Vargas, A.C. (2013) 'Diversity of seM in Streptococcus equi subsp. equi isolated from strangles outbreaks', *Vet Microbiol*, 162(2-4), pp. 663-9.
- Liu, J., Wang, Y., Li, Z., Ren, Y., Zhao, Y. and Zhao, G. (2018) 'Efficient production of high-molecular-weight hyaluronic acid with a two-stage fermentation', *RSC Advances*, 8(63), pp. 36167-36171.
- Liu, L., Liu, Y., Li, J., Du, G. and Chen, J. (2011) 'Microbial production of hyaluronic acid: current state, challenges, and perspectives', *Microb Cell Fact*, 10, p. 99.
- Liu, L., Wang, M., Du, G. and Chen, J. (2008) 'Enhanced hyaluronic acid production of Streptococcus zooepidemicus by an intermittent alkaline-stress strategy', *Lett Appl Microbiol*, 46(3), pp. 383-8.
- Lu, C.L., Liu, C.Y., Huang, Y.T., Liao, C.H., Teng, L.J., Turnidge, J.D. and Hsueh, P.R. (2011) 'Antimicrobial susceptibilities of commonly encountered bacterial isolates to fosfomycin determined by agar dilution and disk diffusion methods', *Antimicrob Agents Chemother*, 55(9), pp. 4295-301.
- Lu, J., Zhu, Y., Sun, H., Liang, S., Leng, F. and Li, H. (2016) 'Highly efficient production of hyaluronic acid by S. zooepidemicus R42 derived from heterologous expression of bacterial hemoglobin and mutant selection', *Lett Appl Microbiol*.
- Lukacova, M., Barak, I. and Kazar, J. (2008) 'Role of structural variations of polysaccharide antigens in the pathogenicity of Gram-negative bacteria', *Clin Microbiol Infect*, 14(3), pp. 200-6.
- Marcellin, E., Steen, J.A. and Nielsen, L.K. (2014) 'Insight into hyaluronic acid molecular weight control', *Appl Microbiol Biotechnol*, 98(16), pp. 6947-56.

Marcellin, E., Gruber, C.W., Archer, C., Craik, D.J. and Nielsen, L.K. (2009) 'Proteome analysis of the hyaluronic acid-producing bacterium, *Streptococcus zooepidemicus*', *Proteome Sci*, 7, p. 13.

Marcellin, E., Chen, W.Y. and Nielsen, L.K. (2010) 'Understanding plasmid effect on hyaluronic acid molecular weight produced by *Streptococcus equi* subsp. *zooepidemicus*', *Metab Eng*, 12(1), pp. 62-9.

McIver, K.S. and Myles, R.L. (2002) 'Two DNA-binding domains of Mga are required for virulence gene activation in the group A streptococcus', *Molecular Microbiology*, 43, pp. 1591-1601.

McIver, K.S., Thurman, A.S. and Scott, J.R. (1999) 'Regulation of *mga* Transcription in the Group A *Streptococcus*: Specific Binding of Mga within Its Own Promoter and Evidence for a Negative Regulator', *Journal of Bacteriology*, 181(17), pp. 5373-5383.

Menzel, E.J. and Farr, C. (1998) 'Hyaluronidase and its substrate hyaluronan: biochemistry, biological activities and therapeutic uses', *Cancer Letters*, 131, pp. 3-11.

Merant, C., Sheoran, A. and Timoney, J.F. (2011) 'Association of *Streptococcus equi* with equine monocytes', *Veterinary immunology and immunopathology*, 143(1-2), pp. 83-6.

Metzgar, D. and Zampolli, A. (2011) 'The M protein of group A *Streptococcus* is a key virulence factor and a clinically relevant strain identification marker', *Virulence*, 2(5), pp. 402-12.

Michalopoulos, A.S., Livaditis, I.G. and Gougoutas, V. (2011) 'The revival of fosfomicin', *Int J Infect Dis*, 15(11), pp. e732-9.

Misra, S., Hascall, V.C., Markwald, R.R. and Ghatak, S. (2015) 'Interactions between Hyaluronan and Its Receptors (CD44, RHAMM) Regulate the Activities of Inflammation and Cancer', *Front Immunol*, 6, p. 201.

Molina-López, J., Sanschagrin, F. and Levesque, R.C. (2006) 'A peptide inhibitor of MurA UDP-N-acetylglucosamine enolpyruvyl transferase: the first committed step in peptidoglycan biosynthesis', *Peptides*, 27(12), pp. 3115-21.

Monslow, J., Govindaraju, P. and Pure, E. (2015) 'Hyaluronan - a functional and structural sweet spot in the tissue microenvironment', *Front Immunol*, 6, p. 231.

Moreira, M., Noschang, J., Neiva, I.F., Carvalho, Y., Higuti, L.H. and Vicente, V.A. (2010) 'Methodological Variations in the Isolation of Genomic DNA from *Streptococcus* Bacteria', *Brazilian Archives of Biology and Technology*, 53(4), pp. 845-849.

- Moscovici, M. (2015) 'Present and future medical applications of microbial exopolysaccharides', *Front Microbiol*, 6, p. 1012.
- Musatov, A. and Sedláč, E. (2017) 'Role of cardiolipin in stability of integral membrane proteins', *Biochimie*, 142, pp. 102-111.
- Muto, J., Sayama, K., Gallo, R.L. and Kimata, K. (2019) 'Emerging evidence for the essential role of hyaluronan in cutaneous biology', *J Dermatol Sci*, 94(1), pp. 190-195.
- Mylvaganam, H., Bjorvatn, B., Hofstad, T. and Osland, A. (2000) 'Molecular characterization and allelic distribution of the phage-mediated hyaluronidase genes hylP and hylP2 among group A streptococci from western Norway', *Microb Pathog*, 29(3), pp. 145-53.
- Nannini, E.C., Teng, F., Singh, K.V. and Murray, B.E. (2005) 'Decreased virulence of a gls24 mutant of *Enterococcus faecalis* OG1RF in an experimental endocarditis model', *Infect Immun*, 73(11), pp. 7772-4.
- Necas, J., Bartosikova, L., Brauner, P. and Kolar, J. (2008) 'Hyaluornic acid (hyaluronan): a review', *Veterinarni Medicina*, 8, pp. 397-411.
- Nikbakht, M., Karbasi, S., Rezayat, S.M., Tavakol, S. and Sharifi, E. (2019) 'Evaluation of the effects of hyaluronic acid on poly (3-hydroxybutyrate)/chitosan/carbon nanotubes electrospun scaffold: structure and mechanical properties', *Polymer-Plastics Technology and Materials*, 58(18), pp. 2031-2040.
- Nizet, V. (2007) 'Understanding how leading bacterial pathogens subvert innate immunity to reveal novel therapeutic targets', *J ALLERGY CLIN IMMUNOL*, pp. 13-22.
- North, S.E., Wakeley, P.R., Mayo, N., Mayers, J. and Sawyer, J. (2014) 'Development of a real-time PCR to detect *Streptococcus equi* subspecies *equi*', *Equine Vet J*, 46(1), pp. 56-9.
- Novick, R.P. (2000) 'Sortase: the surface protein anchoring transpeptidase and the LPXTG motif', *Trends in Microbiology*, 8(4).
- Nwodo, U.U., Green, E. and Okoh, A.I. (2012) 'Bacterial exopolysaccharides: functionality and prospects', *Int J Mol Sci*, 13(11), pp. 14002-15.
- Oliveira, A.H., Ogrodowski, C.S., Macedo, A.C., Santana, M.H. and Goncalves, L.R.B. (2013) 'Cashew apple juice as microbial cultivation medium for non-immunogenic hyaluronic acid production', *Brazilian Journal of Microbiology*, 4, pp. 1097-1104.

Opdyke, J.A., Scott, J.R. and Morgan Jr, C.P. (2001) 'A secondary RNA polymerase sigma factor from *Streptococcus pyogenes*', *Molecular Microbiology*, 42(2), pp. 495-502.

Paget, M.S. and Helmann, J.D. (2003) 'The σ^{70} family of sigma factors', *Genome Biology*, 4(203).

Papadimitriou, K., Alegria, A., Bron, P.A., de Angelis, M., Gobbetti, M., Kleerebezem, M., Lemos, J.A., Linares, D.M., Ross, P., Stanton, C., Turrone, F., van Sinderen, D., Varmanen, P., Ventura, M., Zuniga, M., Tsakalidou, E. and Kok, J. (2016) 'Stress Physiology of Lactic Acid Bacteria', *Microbiol Mol Biol Rev*, 80(3), pp. 837-90.

Patil, K.P., Patil, D.K., Chaudhari, B.L. and Chincholkar, S.B. (2011) 'Production of hyaluronic acid from *Streptococcus zooepidemicus* MTCC 3523 and its wound healing activity', *J Biosci Bioeng*, 111(3), pp. 286-8.

Pelkonen, S., Lindahl, S.B., Suomala, P., Karhukorpi, J., Vuorinen, S., Koivula, I., Vaisanen, T., Pentikainen, J., Autio, T. and Tuuminen, T. (2013) 'Transmission of *Streptococcus equi* subspecies *zooepidemicus* infection from horses to humans', *Emerg Infect Dis*, 19(7), pp. 1041-8.

Pires, A.M., Eguchi, S.Y. and Santana, M.H. (2010) 'The influence of mineral ions on the microbial production and molecular weight of hyaluronic acid', *Appl Biochem Biotechnol*, 162(8), pp. 2125-35.

Pires, A.M. and Santana, M.H. (2010) 'Metabolic effects of the initial glucose concentration on microbial production of hyaluronic acid', *Appl Biochem Biotechnol*, 162(6), pp. 1751-61.

Pires, A.M., Macedo, A.C., Eguchi, S.Y. and Santana, M.H. (2010) 'Microbial production of hyaluronic acid from agricultural resource derivatives', *Bioresour Technol*, 101(16), pp. 6506-9.

Pontes-Quero, G.M., Garcia-Fernandez, L., Aguilar, M.R., San Roman, J., Perez Cano, J. and Vazquez-Lasa, B. (2019) 'Active viscosupplements for osteoarthritis treatment', *Semin Arthritis Rheum*, 49(2), pp. 171-183.

Poyart, C., Pellegrini, E., Gaillot, O., Boumaila, C., Baptista, M. and Trieu-Cuot, P. (2001) 'Contribution of Mn-cofactored superoxide dismutase (SodA) to the virulence of *Streptococcus agalactiae*', *Infect Immun*, 69(8), pp. 5098-106.

Prasad, S.B., Ramachandran, K.B. and Jayaraman, G. (2012) 'Transcription analysis of hyaluronan biosynthesis genes in *Streptococcus zooepidemicus* and metabolically engineered *Lactococcus lactis*', *Appl Microbiol Biotechnol*, 94(6), pp. 1593-607.

Prediger, E. (2012) *Starting with RNA—one-step or two-step RTqPCR?* Available at: <https://www.idtdna.com/pages/education/decoded/article/one-step-two-step> (Accessed: 29 August 2019).

Price, Z.K., Lokman, N.A. and Ricciardelli, C. (2018) 'Differing Roles of Hyaluronan Molecular Weight on Cancer Cell Behavior and Chemotherapy Resistance', *Cancers (Basel)*, 10(12).

Pusterla, N., Leutenegger, C.M., Barnum, S.M. and Byrne, B.A. (2018) 'Use of quantitative real-time PCR to determine viability of *Streptococcus equi* subspecies *equi* in respiratory secretions from horses with strangles', *Equine Vet J*, 50(5), pp. 697-700.

Qiao, Y., Leng, C., Liu, G., Zhang, Y., Lv, X., Chen, H., Sun, J. and Feng, Z. (2019) 'Transcriptomic and proteomic profiling revealed global changes in *Streptococcus thermophilus* during pH-controlled batch fermentations', *Journal of Microbiology*, 57(9), pp. 769-780.

Rehm, B.H. (2010) 'Bacterial polymers: biosynthesis, modifications and applications', *Nat Rev Microbiol*, 8(8), pp. 578-92.

Salwowska, N.M., Bebenek, K.A., Zazdło, D.A. and Weisło-Dziadecka, D.L. (2016) 'Physiochemical properties and application of hyaluronic acid: a systematic review', *Journal of Cosmetic Dermatology*, 15, pp. 520-526.

Samland, A.K., Amrhein, N. and Macheroux, P. (1999) 'Lysine 22 in UDP-N-Acetylglucosamine Enolpyruvyl Transferase from *Enterobacter cloacae* Is Crucial for Enzymatic Activity and the Formation of Covalent Adducts with the Substrate Phosphoenolpyruvate and the Antibiotic Fosfomycin', *Biochemistry*, 38, pp. 13162-13169.

Schommer, N.N., Muto, J., Nizet, V. and Gallo, R.L. (2014) 'Hyaluronan breakdown contributes to immune defense against group A *Streptococcus*', *J Biol Chem*, 289(39), pp. 26914-21.

Schrager, H.M., Rheinwald, J.G. and Wessels, M.R. (1996) 'Hyaluronic acid capsule and the role of streptococcal entry into keratinocytes in invasive skin infection', *J Clin Invest*, 98(9), pp. 1954-8.

Scott, J.R. and Barnett, T.C. (2006) 'Surface proteins of gram-positive bacteria and how they get there', *Annu Rev Microbiol*, 60, pp. 397-423.

Segura, M. and Gottschalk, M. (2005) 'Extracellular Virulence Factors of *Streptococci* Associated With Animal Diseases', *Frontiers in Bioscience*, 9, pp. 1157-1188.

Shahin, K., Thompson, K.D., Inglis, N.F., McLean, K., Ramirez-Paredes, J.G., Monaghan, S.J., Hoare, R., Fontaine, M., Metselaar, M. and Adams, A. (2018) 'Characterization of the outer membrane proteome of *Francisella noatunensis* subsp. *orientalis*', *J Appl Microbiol*, 125(3), pp. 686-699.

Sheldon, W.L., Macauley, M.S., Taylor, E.J., Robinson, C.E., Charnock, S.J., Davies, G.J., Vocadlo, D.J. and Black, G.W. (2006) 'Functional analysis of a group A streptococcal glycoside hydrolase Spy1600 from family 84 reveals it is a beta-N-acetylglucosaminidase and not a hyaluronidase', *Biochem J*, 399(2), pp. 241-7.

Sheng, J.Z., Ling, P.X., Zhu, X.Q., Guo, X.P., Zhang, T.M., He, Y.L. and Wang, F.S. (2009) 'Use of induction promoters to regulate hyaluronan synthase and UDP-glucose-6-dehydrogenase of *Streptococcus zooepidemicus* expression in *Lactococcus lactis*: a case study of the regulation mechanism of hyaluronic acid polymer', *J Appl Microbiol*, 107(1), pp. 136-44.

Sheriff, M.A. and Mahendraku, M. (2015) 'Whole Cell Protein Profiling of Human Pathogenic Bacteria Isolated from Clinical Samples', *Asian Journal of Scientific Research*, 8(3), pp. 374-380.

Shevchenko, A., Jensen, O., Podtelejnikov, A., Sagliocco, F., Wilm, M., Vorm, O., Mortensen, P., Shevchenko, A., Boucherie, H. and Mann, M. (1996) 'Linking genome and proteome by mass spectrometry: Large-scale identification of yeast proteins from two dimensional gels', *Proceedings of the National Academy of Sciences of the United States of America*, 25, pp. 14440-14445.

Silhavy, T.J., Kahne, D. and Walker, S. (2010) 'The Bacterial Cell Envelope', *Cold Spring Harbor Perspectives in Biology*, 2(5), pp. a000414-a000414.

Silver, L.L. (2017) 'Fosfomycin: Mechanism and Resistance', *Cold Spring Harb Perspect Med*, 7(2).

Singhal, N., Maurya, A.K. and Viridi, J.S. (2019) 'Bacterial Whole Cell Protein Profiling: Methodology, Applications and Constraints', *Current Proteomics*, 16(2), pp. 102-109.

Sinner, S.W. and Tunkel, A.R. (2015) 'Viridans Streptococci, Nutritionally Variant Streptococci, Groups C and G Streptococci, and Other Related Organisms', in Bennett, J. E., Dolin, R. and Blaser, M. J. (eds.) *Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases (Eighth Edition)* (2). Eighth edn., pp. 2349-2361.e2.

Slauch, J.M. (2011) 'How does the oxidative burst of macrophages kill bacteria? Still an open question', *Mol Microbiol*, 80(3), pp. 580-3.

Smeesters, P.R., McMillan, D.J. and Sriprakash, K.S. (2010) 'The streptococcal M protein: a highly versatile molecule', *Trends Microbiol*, 18(6), pp. 275-82.

Solis, M.A., Chen, Y.H., Wong, T.Y., Bittencourt, V.Z., Lin, Y.C. and Huang, L.L. (2012) 'Hyaluronan regulates cell behavior: a potential niche matrix for stem cells', *Biochem Res Int*, 2012, p. 346972.

Solish, N., Bertucci, V., Percec, I., Wagner, T., Nogueira, A. and Mashburn, J. (2019) 'Dynamics of hyaluronic acid fillers formulated to maintain natural facial expression', *J Cosmet Dermatol*, 18(3), pp. 738-746.

Somkuti, G.A. and Gilbreth, S.E. (2007) 'Influence of organic buffers on bacteriocin production by *Streptococcus thermophilus* ST110', *Curr Microbiol*, 55(2), pp. 173-7.

Speck, S., Honer, O.P., Wachter, B. and Fickel, J. (2008) 'Characterization of *Streptococcus equi* subsp. *ruminatorum* isolated from spotted hyenas (*Crocuta crocuta*) and plains zebras (*Equus burchelli*), and identification of a M-like protein (SrM) encoding gene', *Vet Microbiol*, 128(1-2), pp. 148-59.

Staerck, C., Gastebois, A., Vandeputte, P., Calenda, A., Larcher, G., Gillmann, L., Papon, N., Bouchara, J.P. and Fleury, M.J.J. (2017) 'Microbial antioxidant defense enzymes', *Microb Pathog*, 110, pp. 56-65.

Starr, C.R. and Engleberg, N.C. (2006) 'Role of hyaluronidase in subcutaneous spread and growth of group A streptococcus', *Infect Immun*, 74(1), pp. 40-8.

Steward, K.F., Robinson, C. and Waller, A.S. (2016) 'Transcriptional changes are involved in phenotype switching in *Streptococcus equi* subspecies *equi*', *Mol Biosyst*, 12(4), pp. 1194-200.

Steward, K.F., Robinson, C., Holden, M.T.G., Harris, S.R., Ros, A.F., Perez, G.C., Baselga, R. and Waller, A.S. (2017) 'Diversity of *Streptococcus equi* subsp. *zooepidemicus* strains isolated from the Spanish sheep and goat population and the identification, function and prevalence of a novel arbutin utilisation system', *Vet Microbiol*, 207, pp. 231-238.

Stollerman, G.H. and Dale, J.B. (2008) 'The importance of the group a streptococcus capsule in the pathogenesis of human infections: a historical perspective', *Clin Infect Dis*, 46(7), pp. 1038-45.

Stoyanov, J.V., Mancini, S., Lu, Z.H., Mourlane, F., Poulsen, K.R., Wimmer, R. and Solioz, M. (2010) 'The stress response protein Gls24 is induced by copper and interacts with the CopZ copper chaperone of *Enterococcus hirae*', *FEMS Microbiol Lett*, 302(1), pp. 69-75.

Sun, X., Wang, Z., Bi, Y., Wang, Y. and Liu, H. (2015) 'Genetic and functional characterization of the hyaluronate lyase HylB and the beta-N-acetylglucosaminidase HylZ in *Streptococcus zooepidemicus*', *Curr Microbiol*, 70(1), pp. 35-42.

Svensater, G., Sjogreen, B. and Hamilton, I.R. (2000) 'Multiple stress responses in *Streptococcus mutans* and the induction of general and stress-specific proteins', *Microbiology*, 146, pp. 107-117.

Sze, J.H., Brownlie, J.C. and Love, C.A. (2016) 'Biotechnological production of hyaluronic acid: a mini review', 3 *Biotech*, 6(1), p. 67.

Takara (2018) 'One-step vs. two-step RT-qPCR—tips for choosing the right protocol', *Tips and Troubleshooting*. Available at: <https://www.takarabio.com/about/bioview-blog/tips-and-troubleshooting/one-step-vs-two-step-rt-qpcr>.

Tamer, T.M. (2013) 'Hyaluronan and synovial joint: function, distribution and healing', *Interdiscip Toxicol*, 6(3), pp. 111-25.

Teng, S., Srivastava, A.K., Schwartz, C.E., Alexov, E. and Wang, L. (2010) 'Structural assessment of the effects of amino acid substitutions on protein stability and protein protein interaction', *Int J Comput Biol Drug Des*, 3(4), pp. 334-49.

Tian, X., Azpurua, J., Hine, C., Vaidya, A., Myakishev-Rempel, M., Abulaeva, J., Mao, Z., Nevo, E., Gorbunova, V. and Seluanov, A. (2013) 'High-molecular-mass hyaluronan mediates the cancer resistance of the naked mole rat', *Nature*, 499(7458), pp. 346-9.

Timoney, J.F. (2004) 'The pathogenic equine streptococci', *Vet Res*, 35(4), pp. 397-409.

Timoney, J.F. (2010) 'Streptococcus'.

Timoney, J.F., Suther, P., Velineni, S. and Artiushin, S.C. (2014) 'The Antiphagocytic Activity of SeM of *Streptococcus equi* Requires Capsule', *Journal of Equine Science*, 25(2), pp. 53-56.

Timoney, J.F., Artiushin, S.C. and Boschwitz, J.S. (1997) 'Comparison of the Sequences and Functions of *Streptococcus equi* M-Like Proteins SeM and SzPSe', *Infection and Immunity*, 65(9), pp. 3600-3605.

Tiwari, S., Patil, R. and Bahadur, P. (2018) 'Polysaccharide Based Scaffolds for Soft Tissue Engineering Applications', *Polymers (Basel)*, 11(1).

Tlapak-Simmons, V.L., Kempner, E.S., Baggenstoss, B.A. and Weigel, P.H. (1998) 'The Active Streptococcal Hyaluronan Synthases (HASs) Contain a Single HAS Monomer and Multiple Cardiolipin Molecules', *The Journal of Biological Chemistry*, 273, pp. 26100-26109.

- Treangen, T.J. and Salzberg, S.L. (2011) 'Repetitive DNA and next-generation sequencing: computational challenges and solutions', *Nature Reviews Genetics*, 13(1), pp. 36-46.
- Tremblay, Y.D., Lo, H., Li, Y.H., Halperin, S.A. and Lee, S.F. (2009) 'Expression of the *Streptococcus mutans* essential two-component regulatory system VicRK is pH and growth-phase dependent and controlled by the LiaFSR three-component regulatory system', *Microbiology*, 155(Pt 9), pp. 2856-65.
- Tripodo, G., Trapani, A., Torre, M.L., Giammona, G., Trapani, G. and Mandracchia, D. (2015) 'Hyaluronic acid and its derivatives in drug delivery and imaging: Recent advances and challenges', *Eur J Pharm Biopharm*, 97(Pt B), pp. 400-16.
- Trombino, S., Servidio, C., Curcio, F. and Cassano, R. (2019) 'Strategies for Hyaluronic Acid-Based Hydrogel Design in Drug Delivery', *Pharmaceutics*, 11(8).
- Tsepilov, R.N. and Beloded, A.V. (2015) 'Hyaluronic Acid--an "Old" Molecule with "New" Functions: Biosynthesis and Depolymerization of Hyaluronic Acid in Bacteria and Vertebrate Tissues Including during Carcinogenesis', *Biochemistry (Mosc)*, 80(9), pp. 1093-108.
- Uchiyama, S., Sun, J., Fukahori, K., Ando, N., Wu, M., Schwarz, F., Siddiqui, S.S., Varki, A., Marth, J.D. and Nizet, V. (2019) 'Dual actions of group B *Streptococcus* capsular sialic acid provide resistance to platelet-mediated antimicrobial killing', *Proc Natl Acad Sci U S A*, 116(15), pp. 7465-7470.
- Vazquez, J.A., Pastrana, L., Pineiro, C., Teixeira, J.A., Perez-Martin, R.I. and Amado, I.R. (2015) 'Production of Hyaluronic Acid by *Streptococcus zooepidemicus* on Protein Substrates Obtained from *Scyliorhinus canicula* Discards', *Mar Drugs*, 13(10), pp. 6537-49.
- Velineni, S. and Timoney, J.F. (2018) 'Evidence for involvement of the Fas BCAX regulon in capsule synthesis by *Streptococcus equi*', *Veterinary Microbiology*, 219, pp. 113-116.
- Velineni, S. and Timoney, J.F. (2015) 'Capsular hyaluronic acid of equine isolates of *Streptococcus zooepidemicus* is upregulated at temperatures below 35 degrees C', *Equine Vet J*, 47(3), pp. 333-8.
- Wacker, M.J. and Godard, M.P. (2005) 'Analysis of One-Step and Two-Step Real-Time RT-PCR Using SuperScript III', *Journal of Molecular Techniques*, 16, pp. 266-271.
- Wajima, T., Morozumi, M., Hanada, S., Sunaoshi, K., Chiba, N., Iwata, S. and Ubukata, K. (2016) 'Molecular Characterization of Invasive *Streptococcus dysgalactiae* subsp. *equisimilis*, Japan', *Emerg Infect Dis*, 22(2), pp. 247-54.

Waller, A.S., Paillot, R. and Timoney, J.F. (2011) 'Streptococcus equi: a pathogen restricted to one host', *J Med Microbiol*, 60(Pt 9), pp. 1231-40.

Wang, Y., Yi, L., Wu, Z., Shao, J., Liu, G., Fan, H., Zhang, W. and Lu, C. (2012) 'Comparative proteomic analysis of Streptococcus suis biofilms and planktonic cells that identified biofilm infection-related immunogenic proteins', *PLoS One*, 7(4), p. e33371.

Watson, E., Alberdi, M.P., Inglis, N.F., Lainson, A., Porter, M.E., Manson, E., Imrie, L., McLean, K. and Smith, D.G.E. (2014a) 'Proteomic analysis of Lawsonia intracellularis reveals expression of outer membrane proteins during infection', *Vet Microbiol*, 174(3-4), pp. 448-455.

Watson, E., Sherry, A., Inglis, N.F., Lainson, A., Jyothi, D., Yaga, R., Manson, E., Imrie, L., Everest, P. and Smith, D.G. (2014b) 'Proteomic and genomic analysis reveals novel Campylobacter jejuni outer membrane proteins and potential heterogeneity', *EuPA Open Proteom*, 4, pp. 184-194.

Webb, K., Barker, C., Harrison, T., Heather, Z., Steward, K.F., Robinson, C., Newton, J.R. and Waller, A.S. (2013) 'Detection of Streptococcus equi subspecies equi using a triplex qPCR assay', *Vet J*, 195(3), pp. 300-4.

Weigel, P.H. and Baggenstoss, B.A. (2012) 'Hyaluronan synthase polymerizing activity and control of product size are discrete enzyme functions that can be uncoupled by mutagenesis of conserved cysteines', *Glycobiology*, 22(10), pp. 1302-10.

Weigel, P.H. (2015) 'Hyaluronan Synthase: The Mechanism of Initiation at the Reducing End and a Pendulum Model for Polysaccharide Translocation to the Cell Exterior', *Int J Cell Biol*, 2015, p. 367579.

Weigel, P.H. (2002) 'Functional Characteristics and Catalytic Mechanisms of the Bacterial Hyaluronan Synthases', *Life*, 54, pp. 201-211.

Weigel, P.H. and DeAngelis, P.L. (2007) 'Hyaluronan synthases: a decade-plus of novel glycosyltransferases', *J Biol Chem*, 282(51), pp. 36777-81.

Wessels, M.R., Moses, A.E., Goldberg, J.B. and DiCesare, T.J. (1991) 'Hyaluronic acid capsule is a virulence factor for mucoid group A Streptococci', *Proc Natl Acad Sci U S A*, 88, pp. 8317-8321.

Wong, S.S. and Yuen, K.Y. (2012) 'Streptococcus pyogenes and re-emergence of scarlet fever as a public health problem', *Emerg Microbes Infect*, 1(7), p. e2.

Wu, T.F., Huang, W.C., Chen, Y.C., Tsay, Y.G. and Chang, C.S. (2009) 'Proteomic investigation of the impact of oxygen on the protein profiles of hyaluronic acid-producing *Streptococcus zooepidemicus*', *Proteomics*, 9(19), pp. 4507-18.

Yamada, T. and Kawasaki, T. (2005) 'Microbial synthesis of hyaluronan and chitin: New approaches', *J Biosci Bioeng*, 99(6), pp. 521-8.

Yokogawa, K., Kawata, S., Nishimura, S., Ikeda, Y. and Yoshimura, Y. (1974) 'Mutanolysin, Bacteriolytic Agent for Cariogenic Streptococci: Partial Purification and Properties', *Antimicrob Agents Chemother*, 6(2), pp. 156-165.

Yu, H. and Stephanopoulos, G. (2008) 'Metabolic engineering of *Escherichia coli* for biosynthesis of hyaluronic acid', *Metab Eng*, 10(1), pp. 24-32.

Zakeri, A. and Rasaee, M. (2016) 'Identification of Wild Type *Streptococcus Zooepidemicus* and Optimization of Culture Medium and Fermentation Conditions for Production of Hyaluronic Acid', *Biosciences, Biotechnology Research Asia*, 13(1), pp. 189-198.

Zhang, J., Ding, X., Yang, L. and Kong, Z. (2006) 'A serum-free medium for colony growth and hyaluronic acid production by *Streptococcus zooepidemicus* NJUST01', *Appl Microbiol Biotechnol*, 72(1), pp. 168-172.

Zhang, L., Huang, H., Wang, H., Chen, J., Du, G. and Kang, Z. (2016) 'Rapid evolution of hyaluronan synthase to improve hyaluronan production and molecular mass in *Bacillus subtilis*', *Biotechnol Lett*.

Zhang, R., Zhou, J., Song, Z. and Huang, Z. (2018) 'Enzymatic properties of beta-N-acetylglucosaminidases', *Appl Microbiol Biotechnol*, 102(1), pp. 93-103.

Zhang, X., Duan, X.-J. and Tan, W.-S. (2010) 'Mechanism for the effect of agitation on the molecular weight of hyaluronic acid produced by *Streptococcus zooepidemicus*', *Food Chemistry*, 119(4), pp. 1643-1646.

Zoeiby, A.E., Sanschagrin, F. and Levesque, R.C. (2003) 'Structure and function of the Mur enzymes: development of novel inhibitors', *Molecular Microbiology*, 47, pp. 1-12.